

# A delegation-based theory of expertise

Attila Ambrus, Volodymyr Baranovskyi and Aaron Kolb \*

## Abstract

We investigate competition in a delegation framework, with a coarsely informed principal. Two imperfectly informed and biased experts simultaneously propose action choices. A principal with a diffuse prior, and only being able to ordinally compare the two proposals, has to choose one of them. The selected expert might receive a bonus payment. We show that having a second expert benefits the principal, even if the two experts have the same biases and the bonus of the winner is zero. In contrast with other models of expertise, in our setting the principal prefers experts with equal rather than opposite biases. Increasing the bonus brings experts closer to truthful reporting, but this only benefits the principal up to a threshold level, with further increases in the bonus strictly decreasing her payoffs. A methodological contribution of our paper is characterizing restrictions on the set of strategies which allows a formal generalization of ex ante expected payoffs to games with diffuse prior.

## 1 Introduction

There are many situations in which a principal lacks the knowledge and expertise to perform a certain task, and therefore has to delegate the job to a qualified expert. Examples include a candidate running for office who has to hire an expert to work out her economic agenda, or the CEO of a pharmaceutical company who must delegate building a research and development division to a scientist. Further complicating the principal's situation is that experts tend to have systemic biases, preferring suboptimal actions from the principal's perspective.

In this paper we investigate a model in which a principal has to delegate a task to one of two imperfectly informed experts. The need to delegate differentiates our model from models of expertise in which experts send cheap talk recommendations to the principal, such as [Krishna and Morgan \(2001b\)](#). In particular, we consider the following game. First experts receive noisy and conditionally independent signals of a single dimensional state variable. The principal's ideal action is equal to the state, but each expert has a constant bias (either positive or negative) and a resulting ideal point different from the sender's. Next, the experts simultaneously propose actions. A proposal is assumed to bind the expert to perform the given action whenever the principal delegates the task to him.<sup>1</sup> The principal then chooses one of the two offers, and the corresponding action is taken by the given expert. We also allow for the possibility that the experts care about being selected, on top of caring about the implemented action (which is independent of who takes the action). In particular,

---

\*We thank Peter Landry, Xiao Yu Wang, Navin Kartik, Ricardo Alonso, David McAdams, Curtis Taylor and Scott Kominers for useful feedback. All errors are our own.

<sup>1</sup>Even if the principal might not have the knowledge to verify whether the expert indeed chose the action that he proposed, outside experts might be able to verify if that was the case and hence penalties can be imposed on experts deviating from their proposals.

we allow for a bonus to the chosen expert, either as a monetary payment or as a non-monetary benefit (such as increased prestige in the profession). We investigate both the case when the bonus is exogenously given, and when it is optimally chosen by the principal.

Our model best applies to situations in which the principal lacks the knowledge to implement or initiate changes in the proposed actions, so all she can do is solicit different proposals and choose one of them. The assumption that proposals commit the experts correspond to common law, according to which an offer is a statement of terms on which the offeror is willing to be bound, and it shall become binding as soon as it is accepted by the person to whom it is addressed.<sup>2</sup> A different application for our model is political competition: starting from [Downs \(1957\)](#), most papers on political competition in a [Hotelling \(1929\)](#) framework assume that candidates are committed to the policies they announce in the campaign, and the electorate can only choose between the policies announced by the candidates.<sup>3</sup> In this context the bonus corresponds to the rents from being in office. Yet another application for our model is a setting where a legislative body (floor) seeks legislative proposals for the same bill from multiple committees, using a modified rule (see [Gilligan and Krehbiel \(1989\)](#), [Krishna and Morgan \(2001a\)](#)), meaning that the floor cannot amend the proposals and can only accept one of the proposed bills without modification, conforming to the basic assumptions of our model.<sup>4</sup>

We focus our investigation to situations in which the principal has limited knowledge of the decision environment. Besides imposing the constraint that she cannot perform the task herself, and hence she needs to delegate, we make two additional assumptions along these lines. The first one is that the principal's prior is improper uniform over a state space represented by the real line. The second one is that the principal lacks the ability to measure the difference between two proposed actions, and she can only ordinally compare them. This leaves only a few simple strategies available to the principal: always choose expert 1's offer (effectively delegating the action choice to expert 1), always choose expert 2's offer, always choose the minimum of the two offers, and always choose the maximum of the two offers. This type of coarse information processing is considered in the context of consumer choice problems in [Kamenica \(2008\)](#). The main question we address is how much a principal with such minimal knowledge of the environment can benefit from the presence of multiple experts.

A methodological contribution of our paper is that, as opposed to existing game theoretical models of improper prior ([Friedman \(1991\)](#), [Klemperer \(1999\)](#), [Morris and Shin \(2002, 2003\)](#), [Myatt and Wallace \(2014\)](#)), we provide an alternative to the standard approach of leaving expected payoffs of the game undefined and only interpreting payoffs conditional on signal realizations. In particular, we characterize the maximal set of expert strategies in our game in which the concept of ex ante expected payoffs can be extended in a rigorous sense, despite the improper prior on the state space: strategies for which the limit payoffs of the players for any possible sequence of proper priors diffusing (converging in a formal sense to the improper prior) are always the same. We can then assign these

---

<sup>2</sup>See [Treitel \(1999\)](#), p8.

<sup>3</sup>For theoretical motivations for this assumption, and empirical relevance in the political competition context, see [Pétry and Collette \(2009\)](#), [Kartik et al. \(2015\)](#), and papers cited therein.

<sup>4</sup>[Gilligan and Krehbiel \(1989\)](#) analyze this situation with an additional option to the floor, in the form of not accepting either of the proposals and opting for a status quo outcome. As opposed to our model, [Gilligan and Krehbiel \(1989\)](#) assume perfectly informed experts (committees), which fundamentally changes the strategic interaction.

limit payoffs as ex ante expected payoffs under an improper uniform prior, and as in standard games, have well-defined expected payoffs for every strategy profile. In our game the largest set of expert strategies for which the limit expected payoffs are well-defined, given the set of simple strategies of the principal, are equivalent to strategies in which the difference between the proposed action and the observed signal, which we label as *markup*, is constant over all possible signals.<sup>5</sup> We define the set of strategies for the experts to be such constant markup strategies.

We note that even though we restrict players' strategies in our game to the above simple strategies, in all equilibria we characterize, no player has a unilateral incentive to deviate, at any information set. That is, even if a player was allowed to play any strategy, she would stick to playing a simple strategy as long as the other players play simple strategies. Hence, the equilibria we characterize would remain equilibria in any sensible extension of the game in which players were allowed to choose from larger sets of strategies. In particular, this holds for the game in which the principal's strategy is not restricted to be based on coarse information (the binary relationship between the two proposed actions): the equilibria in our model would remain equilibria with a more sophisticated principal, but particularly simple equilibria in which the principal only uses coarse information.<sup>6</sup>

In our game there always exists a trivial pure strategy equilibrium, in the form of delegating the task to one of the experts. Formally, one expert always proposing his ideal action conditional on the signal he observes (equal to the signal plus his bias), the other expert proposing his signal minus the first expert's bias, and the principal always delegating the task to the first expert constitutes a (Bayesian Nash) equilibrium.<sup>7</sup> The question is whether there exist other equilibria of the game, in which the principal either always chooses the minimum or always chooses the maximum of the two proposals (hence her choice depends nontrivially on the proposals).

Assuming, without loss of generality, that the sum of the biases of the two experts is nonnegative, we show that if the bonus is not too large then there always exists an equilibrium in which on average the experts choose a positive markup and the principal chooses the minimum of the two offers. We call this equilibrium "upward," as the experts on average exaggerate their signals upwards in their proposals. If the experts have the same biases and the bonus is small then in this upward equilibrium both experts propose actions strictly above their ideal actions based purely on their private signals. This is because, similarly to the winner's curse phenomenon in common value auctions, being selected by the principal contains information on the other expert's signal (namely that his signal is higher), changing the optimal action of the expert. In equilibrium, proposals have to be optimal conditional on the event that the other expert's action proposal is higher.

We also show that if the experts' signals are noisy enough then, for a subset of the range of bonuses for which an upward equilibrium exists, there also exists a "downward" equilibrium in which experts propose actions on average *below* their signals, and the maximum of the two proposals is selected by the principal. The strategic forces are similar to those in upward equilibrium: the fact that the maximum of the two offers is selected pushes proposals downwards, and for noisy signals on average

---

<sup>5</sup>More precisely, mixtures of such constant markup strategies.

<sup>6</sup>Of course in the extended game there can be other, more complicated equilibria.

<sup>7</sup>There are other Bayesian Nash equilibria on mixed strategies with the same outcome, in which one expert always proposes his ideal point, the other expert "babbles" (randomizes over possible messages he can send), and the principal always delegates the task to the first expert.

experts modify their proposals downward relative to their signals.<sup>8</sup> However, this type of equilibrium does not exist when the signals are too precise, as in that case the information contained by being selected does not shift the optimal proposals of the experts enough to make markups negative on average.

The feature of the above equilibria that similarly biased experts exaggerate in a particular direction (and hence the principal should choose the proposal least in the direction of the exaggeration) is intuitive, and in line with empirical evidence. For example, [Zitzewitz \(2001\)](#), [Bernhardt et al. \(2006\)](#) and [Chen and Jiang \(2006\)](#) find that financial analysts systematically exaggerate their forecasts relative to unbiased forecasts based on the analysts' information sets, while [Iezzoni et al. \(2012\)](#) report that 55% of doctors in a survey said that in the previous year they had been more positive about patients' prognoses than their medical histories warranted.

We show that even if a downward equilibrium exists, the upward equilibrium always yields a higher payoff to the principal. For this reason, we focus on comparing the principal's expected payoff in the upward equilibrium to the payoff from simple delegation to the less-biased expert. In general the comparison can go either way. However, when the experts have the same biases - as in settings where experts have similar agendas - then the principal is always better off in the upward equilibrium, hence having a second expert can strictly improve the principal's payoff. This is the case even when the bonus is zero, that is when there is no competition among experts for being selected. Indeed, in the latter case the expected bias of the implemented action in the upward equilibrium is equal to the common bias of the experts. Nevertheless, the upward equilibrium yields a higher payoff because it utilizes the private signals of both experts, reducing the variance of the action choice.

We also compare the principal's payoffs when experts have equal versus opposite biases, and our result contrasts sharply with some of the existing literature. In our model, assuming the upward equilibrium is played, having two experts with identical biases yields a higher payoff than having two antagonist experts with opposite biases. In general, the expected bias of the implemented action is smaller with antagonist experts than with experts having the same bias, but this benefit is outweighed by a higher variance of the implemented action that arises because the expert with the lower bias is selected most of the time, and so the information from the other expert's signal is only utilized to a limited extent. This result contrasts models of competition in persuasion ([Milgrom and Roberts \(1986\)](#), [Gentzkow and Kamenica \(2015\)](#)), in which antagonist experts benefit the principal by pressing each other to reveal more information,<sup>9</sup> and with the multi-sender cheap talk model of [Krishna and Morgan \(2001b\)](#), in which having a second sender with the same bias does not benefit the receiver.<sup>10</sup>

The principal's expected payoff depends in a complicated way on the noise in the experts' signals,

---

<sup>8</sup>There can also be an equilibrium in which the principal mixes with a particular probability between accepting the lower or the higher proposal. We provide a partial characterization of such mixed equilibria in the Supplementary Appendix.

<sup>9</sup>Experts with identical agendas can be better for the principal than experts with opposing agendas in the persuasion model of [Bhattacharya and Mukherjee \(2013\)](#). The mechanism is rather different than in our paper, though: with similar experts an undesirable default action can provide strong incentives for both experts to reveal information.

<sup>10</sup>See also [Shin \(1998\)](#) and [Dewatripont and Tirole \(1999\)](#) for different types of models making the case for adversarial procedures.

and on the amount of the bonus. Hence, for these comparative statics we focus on the case of equally biased experts. Even in this case, the effect of the variance of the experts' signals is ambiguous. An increased precision of experts' signals reduces the variance of the implemented action conditional on the state. For small bonuses, this unambiguously increases the principal's expected payoff. However, for larger bonuses, it might benefit the principal in the upward equilibrium if the experts increase their markups,<sup>11</sup> which can result from increasing the variance of the signals. We provide an exact characterization (for equally biased experts) for when a decrease in the variance of experts' signals benefits the principal.

Increasing the bonus reduces the absolute values of the experts' markups, hence bringing their proposals closer to truthful reporting, both in the upward and downward equilibria. Intuitively, a higher bonus increases competition among experts, leading them to decrease their proposals in the upward equilibrium and increase their proposals in the downward equilibrium. In the upward equilibrium, this initially improves the principal's expected payoff by decreasing the expected bias of the implemented action. There is a threshold level of bonus though at which the expected bias of the implemented action becomes zero, and increasing the bonus above this threshold decreases the principal's payoff. When the bonus comes from exogenous sources, the optimal bonus from the principal's perspective is always strictly positive, and is on the interior of the interval of bonuses for which the upward equilibrium exists. When the bonus is paid by the principal, the optimal bonus amount is always strictly smaller than in the previous case, and depending on the parameters it can be either strictly positive or zero.

In the political competition application of the model the result implies that a small amount of office-seeking motivation can be beneficial for voters, but at higher levels a further increase in office-seek motivation can adversely affect voters' welfare.

Lastly, we consider two extensions of our model, for equally biased experts. In the first one we allow the principal to commit ex ante to any mixture of simple strategies, and show that for bonuses that are not too large, such commitment leads to the same outcome as in the upward equilibrium of the original game, hence the ability to commit does not improve the principal's welfare. In the second extension we drop the dependence of the unselected expert's payoff on the implemented action, and instead assume that the expert gets a fixed outside option payoff. This variant of the model is more realistic in market transaction situations, such as when experts are car mechanics or doctors. A car mechanic might be biased towards larger repairs than necessary, but typically he does not care about what type of repair is chosen in case a different mechanic is selected to do the job. The analysis of this version of the model is more involved, but we show that under some parameter restrictions similar upward and downward equilibria exist as in the baseline model. The fact that the unselected expert gets a fixed outside payment increases the experts' proposals in upward equilibrium, and decreases them in downward equilibrium.

The literature on delegation so far mainly focused on either the question of delegating the action

---

<sup>11</sup>This is related to the chunkiness of the principal's possible choices in our model: for certain parameter values sticking with choosing the minimum of two proposed actions is still optimal for the principal, even though it leads to the implemented action being negatively biased. This can happen if choosing the maximum offer would lead to an even larger positive bias. These are the cases when an increase in the expectation of the minimum offer benefits the principal.

choice versus retaining the right to take the action (Dessein (2002), Li and Suen (2004)) or on optimally constraining the action choices of a particular expert (Holmström (1977), Melumad and Shibano (1991), Alonso and Matouschek (2008)). Krishna and Morgan (2008) investigate how monetary incentives can be used optimally in delegation to a single agent. The first paper in a delegation framework with asymmetric information we are aware of that considers competing agents is Kartik et al. (2015), in the context of the classic Downs (1957) model of political competition. Similarly to our setting, politicians receive independent private signals about the state of the world and hence the optimal policy from the electorate’s point of view. The main difference between the models is that the politicians in Kartik et al. (2015) do not have policy preferences, and they are purely office-motivated. For this reason neither the own private information nor the rival’s private information directly affects their expected payoffs, and the candidates play a zero-sum game. In contrast, in our model the experts’ signals are directly payoff-relevant for them, and their interests are partially aligned, as in higher states they would both like to induce higher actions.<sup>12</sup> This leads to different equilibrium dynamics than in Kartik et al. (2015), and to distinct conclusions: in particular, in their model the electorate can never strictly benefit from the presence of a second candidate (relative to just a single one).<sup>13</sup>

Outside the delegation literature, Gerardi et al. (2009) investigates aggregation of expert opinions through a particular mechanism that approximates the first best outcome if signals are very accurate. Pesendorfer and Wolinsky (2003) investigate the effects of being able to solicit a second opinion from a different expert, in a dynamic model in which experts are not biased but it is costly for them to gather information. Another line of literature investigates multi-sender extensions of the cheap talk model of Crawford and Sobel (1982), and finds that under certain conditions there can be equilibria in which the receiver can extract full or almost full information from the senders (Gilligan and Krehbiel (1989), Austen-Smith (1993), Wolinsky (2002), Battaglini (2002, 2004), Ambrus and Takahashi (2008), Ambrus and Lu (2014)). As opposed to the above papers, we investigate settings in which the principal cannot solicit information from experts and then take the action choice herself. Lastly, Ottaviani and Sørensen (2006) consider a model with multiple experts with reputational concerns reporting sequentially on privately observed signals. The issues they focus on (potential herding behavior of experts) are very different than in the current paper.

The coarse information processing on the part of the principal in our model can be considered as a particular instance of analogy classes, proposed in Jehiel (2005) and Jehiel and Koessler (2008).<sup>14</sup> In particular, the experts’ proposals can be assigned to three analogy classes: expert 1’s proposed action is higher, expert 2’s proposed action is higher, and the two proposed actions are equal. As

<sup>12</sup>For classic papers in political science literature assuming that politicians have mixed motivation (having both policy preferences and wanting to win), as in our model, see Wittman (1983) and Calvert (1985).

<sup>13</sup>Our model approximates the model in Kartik et al. (2015) when the bonus payment is very large and so the agents mainly care about being selected. We find that for very large bonuses the only equilibria in our model involve delegating the action choice to a single agent, which is in line with the result on maximum informativeness of political competition in Kartik et al. (2015). Correspondingly, Kartik et al. (2015) discuss an extension of their model in which they show that allowing a small amount of ideological motivation for the candidates, and assuming that they are close to unbiased from the electorate’s point of view implies that in equilibrium one candidate must be winning *ex ante* with probability close to 1. These results suggest that there is no discontinuity between no policy preference versus a small amount of policy preference for the agents. Our paper mainly focuses on cases in which agents’ policy preferences are relatively important.

<sup>14</sup>Similar concepts have been introduced in Eyster and Rabin (2005) and in Esponda (2008).

opposed to the examples these papers highlight, the equilibria we characterize in our model would remain equilibria even if the principal could condition her choice on the exact proposals.<sup>15</sup> Hence the outcomes we characterize can also be interpreted as equilibria in which the principal processes information from the proposals in a coarse way, even though she is not restricted to do so.

## 2 The model

We consider the following multi-stage game with incomplete information. There are three players: a principal and two experts. The set of states of the world is  $\mathbb{R}$ , and we assume that the common prior distribution of states is diffuse (improper uniform).

In stage 0 state  $\theta \in \mathbb{R}$  realizes. In stage 1 each expert  $i = 1, 2$  receives a noisy private signal about the state of the world  $s_i = \theta + \epsilon_i$ , where  $\epsilon_i \sim N(0, \sigma^2)$ , and  $\epsilon_1$  and  $\epsilon_2$  are independent. In stage 2 each expert  $i$  proposes an action  $a_i \in \mathbb{R}$  to the principal. In stage 3 the principal chooses one of the two experts, who then implements the action he proposed in stage 2.

Let real-valued functions  $a_1(s_1)$  and  $a_2(s_2)$  denote the strategies of expert 1 and expert 2 respectively, while  $C(a_1, a_2)$  is the principal's choice strategy. If action  $a = a_i$  is taken then the principal receives payoff  $V(a, \theta) = -(a - \theta)^2$ , expert  $i$  receives  $B - (a - \theta - b_i)^2$ , and expert  $j \neq i$  receives  $-(a - \theta - b_j)^2$ . We refer to  $b_i$  as the bias of expert  $i$  and to  $B$  as the bonus received by the chosen expert. Without loss of generality, we assume that  $b_1 + b_2 \geq 0$  and  $b_1 \geq b_2$ . We further assume that all parameters of the game are common knowledge.

We are interested in situations in which the principal is unable to comprehend/measure the exact values of  $a_1$  and  $a_2$ , she can only make a binary comparison and tell which proposed action is higher. Hence, we restrict the principal's strategies to the following simple strategies: (i)  $C(a_1, a_2) = 1$  for any  $a_1, a_2 \in \mathbb{R}$  (always choose expert 1); (ii)  $C(a_1, a_2) = 2$  for any  $a_1, a_2 \in \mathbb{R}$  (always choose expert 2); (iii)  $C(a_1, a_2) = 1$  if  $a_1 < a_2$ ,  $C(a_1, a_2) = 2$  if  $a_2 < a_1$ , and  $C(a_1, a_2) \in \{1, 2\}$  if  $a_1 = a_2$  (always choose the expert with the smaller proposal); and (iv)  $C(a_1, a_2) = 1$  if  $a_1 > a_2$ ,  $C(a_1, a_2) = 2$  if  $a_2 > a_1$ , and  $C(a_1, a_2) \in \{1, 2\}$  if  $a_1 = a_2$  (always choose expert with the higher proposal).

Define a markup of an expert's strategy as a difference between his proposed action and his signal:  $k_i(s_i) = a_i(s_i) - s_i$ . We restrict the experts' strategy sets, to constant markup/markdown strategies:  $a_i(s_i) = s_i + k_i$ , where  $k_1, k_2 \in \mathbb{R}$ . We motivate this restriction in the next section, where we show that given the set of strategies of the expert defined above, constant markup/markdown strategies are essentially the only strategies in our game for which the concept of ex ante expected payoff can be extended to in a formal way. With slight abuse of notation, we use simply  $(k_1, k_2, C(a_1, a_2))$  to denote a strategy profile with constant markup strategies.

We use Bayesian Nash equilibrium (from now on referred to as equilibrium) as our solution concept, and in the main text focus attention on pure strategy equilibria. Given the restrictions we impose on strategies, all the equilibria we characterize in the paper are also perfect Bayesian Nash equilibria. We also note that in all equilibria we characterize, the action choice prescribed by the equilibrium strategy is sequentially rational at any history, even if the restrictions on the

<sup>15</sup>With finer information processing on the part of the principal, there could be other pure strategy equilibria, besides the simple ones characterized in this paper.



set of possible strategies is lifted. That is, the characterized equilibria remain perfect Bayesian Nash equilibria even when strategies are not restricted to be simple (for the principal) and constant markup/markdown (for the experts).

### 3 Microfoundations for Ex Ante Expected Payoffs with a Diffuse Prior

In this section we provide a formal interpretation of a uniform improper prior as a limit construction, and characterize strategies for which ex ante expected payoffs are well-defined in a formal sense.

#### 3.1 Overview

While the uninformative prior is often interpreted as a uniform distribution on the real line, it is not properly defined. We capture the main features of the uninformative prior by using a sequence of (proper) measures that *diffuse* in a formal sense (Definition 3). In this section, we provide two definitions of strategy admissibility and derive a main result for each. We say that a strategy is *weakly admissible* (Definition 4) when ex ante expected payoffs, taken along any diffusing sequence of priors, have a well-defined limit that does not depend on the particular sequence. Similarly, a strategy is *strongly admissible* (Definition 6) if the joint distribution over noise and markups it induces has such a limit. Our first result (Theorem 1) states roughly that in any class of weakly admissible strategies that includes constant markup strategies, every strategy is *nearly stationary* in a particular sense (Definition 5). Our second result (Theorem 2) states that any strongly admissible strategy has a stronger property: it is payoff-equivalent to some mixture of constant markup strategies.

Since we start from strategy sets that are not restricted to be mixtures of constant markup strategies, we need to provide a more careful formal definition of strategies. Since the state space is uncountable, it is not possible to define strategies as products of state-dependent distributions over markups. The key tension is that desirable topologies should be both rich enough so that payoff functions are continuous in strategies, but also coarse enough so that the strategy space is compact. We follow the approach of [Milgrom and Weber \(1985\)](#) in using *distributional strategies*, which are measures  $\mu$  over the product space of signals and markups with bounded support with respect to the latter. A distributional strategy induces a conditional distribution on markups for any given signal, and by integrating over signals, it induces a conditional distribution on actions for any given  $\theta$ .

Below we first provide outlines of the main results detailed in subsections [3.2](#) and [3.3](#).

#### *Weak Admissibility*

Given the form of the experts' loss functions, it is useful to describe strategies in terms of distributions over offers, net of the true state. For every expert strategy  $\mu$  and state  $\theta$ , we define  $G_\mu(\cdot; \theta)$  as this distribution conditional on  $\theta$ . We show that  $G_\mu(\cdot; \theta)$  is continuous in both  $\mu$  and  $\theta$ ,



and the space of all such distributions, denoted  $\mathcal{G}^M$ , is compact (Lemmas 1 and 2).<sup>16</sup> We say that  $\mu$  is *near* a given distribution  $Q \in \mathcal{G}^M$ , and hence *nearly stationary*, if the following property holds: given any  $\eta > 0$ ,  $G_\mu(\cdot, \cdot; \theta)$  is within  $\eta$  distance of  $Q$  for all  $\theta$  except for a set of finite Lebesgue measure (Definition 2).

We begin the proof of Theorem 1 by establishing the existence of some distribution  $Q \in \mathcal{G}^M$  with the following property: for all  $\eta > 0$ ,  $G_\mu(\cdot; \theta)$  is within  $\eta$  of  $Q$  for an infinite measure set of  $\theta$  (Lemma 3). We then call  $Q$  an *attraction*. A necessary condition for  $\mu$  to be near  $Q$  is that  $Q$  is an attraction.

Next, we argue that there can be at most one such attraction for any strategy. The proof of this claim is by contradiction and relies on several steps. We suppose that  $Q$  and  $Q'$  are two distinct attractions. We then show that there must exist some constant markup strategy of the rival,  $\kappa_m$ , against which these distributions yield distinct expected payoffs, denoted  $u_{Q, \kappa_m}, u_{Q', \kappa_m} \in \mathbb{R}$ . Given  $\eta > 0$ , we can construct a sequence of measures  $(\mathbb{P}_n^1)_{n \in \mathbb{N}}$  (resp.  $(\mathbb{P}_n^2)_{n \in \mathbb{N}}$ ) that places increasing mass on  $\theta$  such that  $G_\mu(\cdot; \theta)$  is within  $\eta$  of  $Q$  (resp.  $Q'$ ). By continuity, the limit of expected payoffs must be at most  $\eta$  from  $u_{Q, \kappa_m}$  (resp.  $u_{Q', \kappa_m}$ ). As these are assumed distinct, this is a contradiction.

Finally, given the unique attraction  $Q$ , we show that  $\mu$  is near  $Q$ . If  $\mu$  is near any distribution, that distribution must be an attraction, so  $Q$  is the only candidate. We show that if  $\mu$  is not near  $Q$ , then there is a compact set of distributions that does not contain  $Q$  but contains  $G_\mu(\cdot; \theta)$  for infinitely many  $\theta$ . This compact set itself contains an attraction, and this contradicts the uniqueness of  $Q$ , completing the proof.

### *Strong Admissibility*

For any expert strategy  $\mu$  and state  $\theta$ , we define  $H_\mu(\cdot, \cdot; \theta)$  as the distribution of noise and markups conditional on  $\theta$ . Strong admissibility is essentially the requirement that these distributions have well-defined expectations in the limit (Definition 6), and it implies weak admissibility. As before, we show that  $H_\mu(\cdot, \cdot; \theta)$  is continuous in  $\mu$  and in  $\theta$ , and the space of all such distributions, denoted  $\mathcal{H}^M$ , is compact (Lemmas 4 and 5).

The proof of Theorem 2 begins in the same way as that of Theorem 1. Given a strongly admissible strategy  $\mu$ , we first show the existence of some distribution  $Q \in \mathcal{H}^M$  that is an attraction for  $\mu$ . Next, we argue by contradiction that there can be at most one such attraction for any strategy. Given the unique attraction  $Q$ , we then show that  $\mu$  is near  $Q$ .

To complete the proof, we define  $K_Q$  as the marginal distribution over markups from  $Q$ ; hence  $K_Q$  characterizes a mixture over constant markup strategies. We show that expected payoffs along any diffusing sequence of priors must converge to the same limit under  $\mu$  as under  $K_Q$ , and thus the two strategies are payoff equivalent.

## 3.2 Formal Analysis of Weak Admissibility

We begin by providing a formal definition of an expert's strategy. As mentioned previously, in games of incomplete information with uncountable type spaces, certain topological problems arise

<sup>16</sup>We use the topology of weak convergence for measures and distributions, which is metrized by the Prokhorov distance, and we use the usual topology for  $\mathbb{R}$ ; continuity and compactness are with respect to these topologies.

from interpreting strategies as mixed strategies (i.e., distributions over pure strategies) or behavioral strategies (i.e., products of history-contingent distributions over actions). As an alternative, [Milgrom and Weber \(1985\)](#) introduce the notion of *distributional strategies*, which we adapt in Definition 1.

We let  $S_i = \mathbb{R}$  denote the signal space for expert  $i$ , and  $X_i = \mathbb{R}$  the space of markups. For  $M \in \mathbb{R}$ , we use the notation  $X_i^M := [-M, M]$ . Let  $\lambda$  denote the Lebesgue measure and let  $\gamma$  denote the measure corresponding to  $N(0, \sigma^2)$ . We let  $h(s|\theta)$  denote the PDF of  $N(\theta, \sigma^2)$  and abbreviate  $h(s|0)$  to  $h(s)$ .

**Definition 1.** *A strategy for expert  $i \in \{1, 2\}$  is a probability measure  $\mu_i$  on  $S_i \times X_i$  such that:*

- *for all measurable  $T \subset S_i$ ,  $\mu_i(T \times X_i) = \gamma(T)$ , and*
- *there exists  $M \in \mathbb{R}$  such that  $\text{supp}(\mu(\cdot|s_i)) \subset X_i^M$  for all  $s_i \in S_i$ ,*

*where  $\mu(\cdot|s_i)$  is the regular conditional distribution.*

Since  $S_i \times X_i$  is complete and separable, the space of probability measures  $\Delta(S_i \times X_i)$ , under the topology of weak convergence, is metrized by the Prokhorov distance.<sup>17</sup> We let  $\mathcal{M} \subset \Delta(S_i \times X_i)$  denote the space of all strategies, and  $\mathcal{M}^M := \{\mu \in \mathcal{M} : \text{supp}(\mu(\cdot|s_i)) \subset X_i^M \forall s_i \in S_i\}$ . Note that  $\mathcal{M} = \cup_{M \in \mathbb{R}} \mathcal{M}^M$ .

At face value, a strategy as defined above gives a distribution over signals and markups conditional on  $\theta = 0$ , since the marginal distribution  $\gamma$  is  $N(0, \sigma^2)$ . However, by multiplication of a factor  $\frac{h(s_i|\theta)}{h(s_i|0)}$ , the strategy uniquely determines these distributions conditional on any  $\theta \in \mathbb{R}$ .

For any strategy  $\mu \in \mathcal{M}$  and  $\theta \in \mathbb{R}$  we define a CDF

$$\begin{aligned} G_\mu(x; \theta) &:= \int_{-\infty}^{\infty} \mu((-\infty, x - (s_i - \theta)]|s_i) \frac{h(s_i|\theta)}{h(s_i|0)} \gamma(ds_i) \\ &= \int_{-\infty}^{\infty} \mu((-\infty, x - (s_i - \theta)]|s_i) h(s_i|\theta) \lambda(ds_i). \end{aligned} \tag{1}$$

We interpret  $G_\mu(x; \theta)$  as the joint probability, conditional on  $\theta$ , that the expert makes an offer that exceeds  $\theta$  by at most  $x$ . We define

$$\mathcal{G}^M := \{P \in \Delta(\mathbb{R}) : P \text{ is induced by the CDF } G_\mu(\cdot; \theta) \text{ for some } \theta \in \mathbb{R}, \mu \in \mathcal{M}^M\}.$$

By inspection of (1), it is not difficult to see that for all  $Q \in \mathcal{G}^M$ ,

$$F_{-M} \preceq_{FSD} Q \preceq_{FSD} F_{+M}, \tag{2}$$

where  $\preceq_{FSD}$  denotes first-order stochastic dominance and  $F_k$  denotes the cumulative distribution of  $N(k, \sigma^2)$ . Note also that for any  $\theta \in \mathbb{R}$  and  $\mu \in \mathcal{M}^M$ , there exists  $\mu' \in \mathcal{M}^M$ , defined by  $\mu'(Y|s_i) = \mu(Y|s_i + \theta)$  for all  $Y \subset X^i$  and all  $s \in S_i$ , which satisfies  $G_{\mu'}(\cdot; 0) = G_\mu(\cdot; \theta)$ , and thus  $\mathcal{G}^M$  is the image of the set  $\mathcal{M}^M$  under the map  $G : \mu \mapsto G_\mu(\cdot; 0)$ . Henceforth, we abbreviate  $G_\mu(\cdot; 0)$

---

<sup>17</sup>See [Billingsley \(2009, Theorem 6.8\)](#).

to  $G_\mu$ . We let  $\mathcal{G} := \cup_{M \in \mathbb{R}} G^M$ . As with  $\mathcal{M}^M$  and  $\mathcal{M}$ , we endow  $\mathcal{G}^M$  and  $\mathcal{G}$  with the topology of weak convergence, which is metrized by the Prokhorov distance.

The next two lemmas establish continuity and compactness properties, to be used toward the main result.

**Lemma 1.** *For any  $\mu \in \mathcal{M}$ , the mapping  $G_\mu : \theta \mapsto G_\mu(\cdot; \theta)$  from  $\mathbb{R}$  to  $\mathcal{G}$  is continuous in  $\theta$ .*

**Lemma 2.** *For all  $M \in \mathbb{R}$ ,  $\mathcal{M}^M$  and  $\mathcal{G}^M$  are compact metric spaces.*

**Definition 2.** *For any  $M \in \mathbb{R}$  and any  $\mu \in \mathcal{M}^M$ , we say that  $\mu$  is near a distribution  $Q \in \mathcal{G}^M$  if for all  $\eta > 0$ ,  $\lambda\{\theta : d_P(G_\mu(\cdot; \theta), Q) \geq \eta\} < \infty$ , where  $d_P$  is the Prokhorov metric. We then classify  $\mu$  as nearly stationary.*

It is immediate from the definition that a strategy can be near to at most one distribution in  $\mathcal{G}^M$ . However, a strategy need not be near any distribution; consider, for example, a strategy that assigns markup 1 for all signals  $s \geq 0$  and markup 0 for all signals  $s < 0$ .

Given a strategy profile and bias, we define expert 1's state-dependent expected utility<sup>18</sup>

$$u(\theta; \mu_1, \mu_2, C, b) := - \int_{\mathbb{R}^2} \mathbb{1}\{C(x_1 + \theta, x_2 + \theta) = 1\}(x_1 - b)^2 \\ + \mathbb{1}\{C(x_1 + \theta, x_2 + \theta) = 2\}(x_2 - b)^2 \nu(d(x_1, x_2)),$$

where  $\nu = G_{\mu_1}(\cdot; \theta) \otimes G_{\mu_2}(\cdot; \theta)$  denotes the product measure. When there is no risk of confusion, we write  $u(\theta)$  for brevity.

We formalize the uniform improper prior through *diffusing* sequences of proper measures, defined next. The definition incorporates the intuition that these sequences should increasingly spread mass over the real line and approach uniformity.

**Definition 3.** *Consider a sequence  $(\mathbb{P}_n)_{n \in \mathbb{N}}$  of Borel probability measures on  $\mathbb{R}$ . We say that this sequence is diffusing if for any nonempty and bounded measurable set  $X$  and any  $\eta > 0$ , there exists  $N \in \mathbb{N}$  such that for all  $n \geq N$ ,  $\mathbb{P}_n(X) > 0$  and for any measurable  $Y \subseteq X$ ,  $\left| \frac{\mathbb{P}_n(Y)}{\mathbb{P}_n(X)} - \frac{\lambda(Y)}{\lambda(X)} \right| < \eta$ .*

It follows from this definition that if  $(\mathbb{P}_n)$  is diffusing, then for any bounded and measurable set  $Y$ ,  $\lim_{n \rightarrow \infty} \mathbb{P}_n(Y) = 0$ . To see this, take any such set  $Y$  and any  $\eta > 0$ . Pick any  $M$  such that  $Y \subseteq X := [-M, M]$  and  $\lambda(Y) < 2M\eta = \lambda(X)\eta$ . Then for sufficiently large  $N$ ,  $n \geq N$  implies  $\mathbb{P}_n(X) > 0$  and  $\left| \frac{\mathbb{P}_n(Y)}{\mathbb{P}_n(X)} - \frac{\lambda(Y)}{\lambda(X)} \right| < \eta$ , which implies that  $\mathbb{P}_n(Y) < \frac{\mathbb{P}_n(X)}{\mathbb{P}_n(X)} < \eta + \frac{\lambda(Y)}{\lambda(X)} < 2\eta$ .

For illustrative purposes, we highlight two specific diffusing sequences.

**Example 1.** *Both  $(\mathbb{P}_n^1)$  given by the density  $\frac{1}{2n} \mathbb{1}\{[-n, n]\}$  and  $(\mathbb{P}_n^2)$  given by  $N(0, n)$  are diffusing.*

*Proof.* If  $X$  is bounded and nonempty, then  $X \subset [-n, n]$  for some  $n$ , and then  $\mathbb{P}_n^1(X) = \lambda(X)/(2n) > 0$ . Moreover, if  $Y \subset X$ , then  $\mathbb{P}_n^1(Y) = \lambda(Y)/(2n)$  and thus  $\left| \frac{\mathbb{P}_n^1(Y)}{\mathbb{P}_n^1(X)} - \frac{\lambda(Y)}{\lambda(X)} \right| = 0$ . Now  $\mathbb{P}_n^2(X) > 0$  for all  $n \in \mathbb{N}$ . Let  $h_n$  denote the density of  $N(0, n)$ . Note that for all  $\eta \in (0, 1)$ , for sufficiently large  $n$ ,  $\frac{\inf_{x \in X} h_n(x)}{\sup_{x \in X} h_n(x)} > 1 - \eta$ . Hence  $(1 - \eta)^2 \frac{\lambda(Y)}{\lambda(X)} < \frac{\mathbb{P}_n^2(Y)}{\mathbb{P}_n^2(X)} < (1 - \eta)^{-2} \frac{\lambda(Y)}{\lambda(X)}$ , and by algebra the result follows.  $\square$

<sup>18</sup>It is without loss of generality to focus on expert 1, as the labeling of experts is arbitrary.

Next, we define formally the key property featured in Theorem 1.

**Definition 4.** A class  $\mathcal{M}_0 \subset \mathcal{M}$  of expert strategies is said to be weakly admissible if for any  $\mu_1, \mu_2 \in \mathcal{M}_0$ , principal strategy  $C$ , and  $b \in \mathbb{R}$ :

- For any diffusing sequence  $(\mathbb{P}_n)$  and sufficiently large  $n$ ,  $u(\cdot)$  is  $\mathbb{P}_n$ -integrable and has finite expectation.
- There exists  $u^* \in \mathbb{R}$  such that for any diffusing sequence  $(\mathbb{P}_n)$ ,  $\lim_{n \rightarrow \infty} \int_{\theta \in \mathbb{R}} u(\theta) d\mathbb{P}_n(\theta) = u^*$ .

For any  $m \in \mathbb{R}$ , a constant markup strategy, denoted  $\kappa_m$ , is the unique strategy such that for all  $s_i \in S_i$ ,  $\kappa_m(\{m\}|s_i) = 1$ . Let  $\mathcal{K} := \{\kappa_m : m \in \mathbb{R}\}$  denote the class of constant markup strategies. Note that for constant markup strategies and their mixtures, the function  $u(\cdot)$  defined above is constant in  $\theta$ , and hence the sequence of integrals in Definition 4 trivially converges to the same limit for every diffusing sequence.

**Example 2.** Constant markup strategies and mixtures of constant markup strategies are weakly admissible.

For the proof of the main result, we make use of a weaker property than that of Definition 2.

**Lemma 3.** Let  $(Y, d)$  be a compact metric space,  $X \subseteq \mathbb{R}$  with  $\lambda X = \infty$ , and  $\pi : X \rightarrow Y$  a Lebesgue measurable function. There exists  $y \in Y$  with the property that for all  $\eta > 0$ ,  $\lambda\{x \in X : d(\pi(x), y) < \eta\} = \infty$ . We say that any such  $y$  is an attraction.

We are now ready to state the main result for weak admissibility. The full proof can be found in the appendix.

**Theorem 1.** If  $\mathcal{M}_0$  is weakly admissible and  $\mathcal{K} \subset \mathcal{M}_0$ , then every  $\mu \in \mathcal{M}_0$  is nearly stationary.

### 3.3 Formal Analysis of Strong Admissibility

In order to characterize admissible strategies in a more transparent way, it is useful to define joint distributions over noise and markups given  $\theta$ . For any strategy  $\mu \in \mathcal{M}$  and  $\theta \in \mathbb{R}$  we define a CDF

$$\begin{aligned} H_\mu(\epsilon, x; \theta) &:= \int_{-\infty}^{\epsilon} \mu((-\infty, x]|\hat{\epsilon}_i + \theta)h(\hat{\epsilon}_i + \theta|\theta)\lambda(d\hat{\epsilon}_i) \\ &= \int_{-\infty}^{\epsilon} \mu((-\infty, x]|\hat{\epsilon}_i + \theta)h(\hat{\epsilon}_i)\lambda(d\hat{\epsilon}_i). \end{aligned} \quad (3)$$

The interpretation of  $H_\mu(\epsilon, x; \theta)$  is the probability, conditional on  $\theta$ , that the noise realization is at most  $\epsilon$  and the markup is at most  $x$ . We make several observations analogous to those following (1) in the previous section. Note that for any  $\theta \in \mathbb{R}$  and  $\mu \in \mathcal{M}^M$ , there exists  $\mu' \in \mathcal{M}^M$  with behavior conditional on state 0 matching that of  $\mu$  conditional on state  $\theta$ , and thus in some instances we can assume  $\theta = 0$  without loss of generality. That is, set  $\mu'(Y|s_i) = \mu(Y|s_i + \theta)$  for all measurable  $Y \subset X_i$  and all  $s_i \in S_i$ , which satisfies  $H_{\mu'}(\cdot, \cdot; 0) = H_\mu(\cdot, \cdot; \theta)$ , and thus  $\mathcal{H}^M$ , defined as the image of the set  $\mathcal{M}^M$  under the map  $H : \mu \mapsto H_\mu(\cdot, \cdot; 0)$ , is  $\mathcal{M}^M$  itself. We let  $\mathcal{H} := \cup_{M \in \mathbb{R}} \mathcal{H}^M$ . As with  $\mathcal{M}^M$  and  $\mathcal{M}$ , we endow  $\mathcal{H}^M$  and  $\mathcal{H}$  with the topology of weak convergence and the Prokhorov metric.

**Lemma 4.** For any  $\mu \in \mathcal{M}$ , the mapping  $H_\mu : \theta \mapsto H_\mu(\cdot, \cdot; \theta)$  from  $\mathbb{R}$  to  $\mathcal{H}$  is continuous in  $\theta$ .

The next lemma establishes compactness for the spaces we are analyzing. The compactness of  $\mathcal{M}^M$  is a restatement of Lemma 2. The compactness of  $\mathcal{H}^M = \Delta([-M, M])$  is standard but also follows from Lemma 4 together with the compactness of  $\mathcal{M}^M$ .

**Lemma 5.** For all  $M \in \mathbb{R}$ ,  $\mathcal{M}^M$  and hence  $\mathcal{H}^M$  are compact metric spaces.

The next definition is analogous to Definition 2.

**Definition 5.** For any  $M \in \mathbb{R}$  and any  $\mu \in \mathcal{M}^M$ , we say that  $\mu$  is near a distribution  $Q \in \mathcal{H}^M$  if for all  $\eta > 0$ ,  $\lambda\{\theta : d_P(H_\mu(\cdot, \cdot; \theta), Q) \geq \eta\} < \infty$ , where  $d_P$  is the Prokhorov metric. We then classify  $\mu$  as nearly stationary.

For a strategy to be strongly admissible, we require that it admits a well-defined expected distributional strategy on the product space of noise and markups; that is, the expected joint distribution over noise and markups, when taken along any diffusing sequence, converges to a fixed limiting distribution that is independent of the particular sequence. This condition ensures that for any strategy profile and conditional on any measurable set of signals, expected payoffs are well-defined in that for all diffusing sequences, the vector of expected payoffs converges to the same payoff vector.<sup>19</sup> Our definition uses the fact that convergence in distribution is equivalent to convergence of CDFs at points of continuity (Billingsley, 2009, Theorem 2.1 and Example 2.3).

**Definition 6.** An expert strategy  $\mu \in \mathcal{M}^M$  is said to be strongly admissible if there exists a CDF  $Q$  over  $S_i \times X_i^M$  such that for all  $(\epsilon, x)$  at which  $Q$  is continuous and for all diffusing sequences  $(\mathbb{P}_n)$ ,

$$\lim_{n \rightarrow \infty} \int_{\theta \in \mathbb{R}} H_\mu(\epsilon, x | \theta) d\mathbb{P}_n(\theta) = Q(\epsilon, x).$$

For any  $m \in \mathbb{R}$ , a constant markup strategy, denoted  $\kappa_m$ , is the unique strategy such that for all  $s_i \in S_i$ ,  $\kappa_m(\{m\} | s_i) = 1$ . Let  $\mathcal{K} := \{\kappa_m : m \in \mathbb{R}\}$  denote the class of constant markup strategies. Note that for constant markup strategies, the function  $H_\mu(\epsilon, x | \cdot)$  defined above is constant in  $\theta$ , and hence the sequence of integrals in Definition 6 trivially converges to the same limit for every diffusing sequence.

**Example 3.** All constant markup strategies and their mixtures are strongly admissible.

Given a strongly admissible strategy  $\mu$ , we use Lemma 3 to show the existence of an attraction for  $\mu$ . We show that this attraction, denoted  $Q$ , is unique by way of contradiction; otherwise, there would exist diffusing sequences under which the limiting distributions of Definition 6 differ. Now  $\mu$  is near to at most one strategy (and possibly none). We show that  $\mu$  must be near some strategy,

<sup>19</sup>For any strategy  $\mu$  and state  $\theta$ , let  $P_\mu(y; \theta) := \int_{-\infty}^{\infty} H_\mu(\epsilon, y - \epsilon; \theta) d\epsilon$  denote the probability that an offer exceeds  $\theta$  by at most  $y$ . Strong admissibility guarantees that for any diffusing sequence, these distributions have a well-defined limiting expectation,  $P_\mu^* := \int_{-\infty}^{\infty} Q(\epsilon, y - \epsilon) d\epsilon$ , where  $Q$  is defined as in Definition 6. For any  $\mathbb{P}_n$  and subset of signals, expected payoffs are calculated by integrating (on the outside) with respect to  $\theta$  and (on the inside) with respect to  $y$ . By changing the order of integration and taking limits inside the integral, limiting expected payoffs are thus well-defined as single integrals over  $y$  according to the distribution  $P_\mu^*$ . For payoffs conditional on a subset of signals, the logic is similar.

so it is near  $Q$ , and hence it is nearly stationary. Moreover,  $Q$  defines a mixture of constant markup strategies,  $K_Q$ . Finally, we show that  $\mu$  is payoff equivalent to this mixture in that for any strongly admissible rival strategy, bias, and simple principal strategy, the limiting expected utility under  $\mu$  is the same as under  $Q$ .

**Theorem 2.** *If  $\mu$  is strongly admissible, then it is nearly stationary and payoff equivalent to some mixture of constant markup strategies.*

Motivated by Theorem 2, we hereafter restrict experts' strategies to be (mixtures over) constant markup strategies.

## 4 Equilibria

In this section we characterize all possible pure strategy equilibria in the delegation game. We start with determining the best response of the principal to all possible pairs of constant markup strategies by the experts, and then investigate the candidate pure strategy equilibria consistent with this best response behavior.

### 4.1 Best Response to Constant Markup Strategies

Here we analyze the principal's best response to constant markup strategies, and find that it only depends on the sum of markups. As the principal has a quadratic loss function, the latter can be decomposed into losses from the uncertainty about the true state (which is independent of her action) and the losses from the expected difference between the chosen action and the true state. Therefore the principal tries to choose the offer which is closer to her expectation of the true state. After receiving the signal, the principal's expectation about the true state is lower (higher) than the average of the experts' offers if and only if the sum of the markups is positive (negative). Let  $\arg \min\{a_1, a_2\}$  be defined as  $\{1\}$  if  $a_1 < a_2$ ,  $\{2\}$  if  $a_1 > a_2$ , and  $\{1, 2\}$  if  $a_1 = a_2$ . Similarly, let  $\arg \max\{a_1, a_2\}$  be defined as  $\{1\}$  if  $a_1 > a_2$ ,  $\{2\}$  if  $a_1 < a_2$ , and  $\{1, 2\}$  if  $a_1 = a_2$ .

**Theorem 3.** *If experts follow constant markup strategies  $a_i(s_i) = s_i + k_i$ , then*

- *if  $k_1 + k_2 > 0$ , the principal strictly prefers the lower offer, and  $C(a_1, a_2) \in \arg \min\{a_1, a_2\}$ ;*
- *if  $k_1 + k_2 < 0$ , the principal strictly prefers the higher offer, and  $C(a_1, a_2) \in \arg \max\{a_1, a_2\}$ ;*
- *if  $k_1 + k_2 = 0$ , the principal is indifferent between the offers.*

*Proof.* After observing both offers, the principal updates her belief:  $\theta|a_1, a_2 \sim N(\frac{1}{2}(a_1 + a_2) - \frac{1}{2}(k_1 + k_2), \frac{\sigma^2}{2})$ . Therefore the principal's expected utility from choosing offer  $a_1$  or  $a_2$  is:

$$V(a_1) = \mathbb{E}[-(\theta - a_1)^2] = -\text{Var}(\theta) - (\mathbb{E}[\theta] - a_1)^2 = -\frac{\sigma^2}{2} - \left[\frac{1}{2}(a_2 - a_1) - \frac{1}{2}(k_1 + k_2)\right]^2$$

$$V(a_2) = \mathbb{E}[-(\theta - a_2)^2] = -\text{Var}(\theta) - (\mathbb{E}[\theta] - a_2)^2 = -\frac{\sigma^2}{2} - \left[\frac{1}{2}(a_1 - a_2) - \frac{1}{2}(k_1 + k_2)\right]^2.$$

Hence,  $V(a_1) - V(a_2) = (a_2 - a_1)(k_1 + k_2)$ , which immediately implies the statements in the theorem.  $\square$

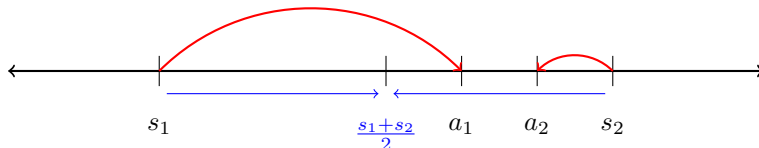


Figure 1:  $k_1 > 0, k_2 < 0, k_1 + k_2 > 0$ . The principal chooses the lower offer  $a_1$ , which lies closer to her expectation  $\frac{s_1 + s_2}{2}$ .

An equilibrium  $(k_1, k_2, C(a_1, a_2))$  is said to be an *upward equilibrium*, if on average experts adjust their signals upwards and the lower proposal is accepted:  $k_1 + k_2 \geq 0$  and  $C(a_1, a_2) \in \arg \min(a_1, a_2)$ . In upward equilibrium, the principal's updated expectation of the state of the world is lower than the average of the two offers. Her best response is to choose the lower offer, which is closer to her expectation, as demonstrated in Figure 1. Likewise, an equilibrium  $(k_1, k_2, C(a_1, a_2))$  is said to be a *downward equilibrium* if  $k_1 + k_2 \leq 0$  and  $C(a_1, a_2) \in \arg \max(a_1, a_2)$ . Note that when  $k_1 + k_2 = 0$  then the principal's posterior expectation of  $\theta$  is exactly the average of the two offers, and she is indifferent between the two. This raises the possibility of equilibria in which the experts play constant markup strategies and the expert mixes between the lower and the higher offer with some fixed probability. We investigate such equilibria in the Supplementary Appendix.

## 4.2 Simple delegation

In our game there always exist simple pure strategy equilibria in which the principal always chooses the same expert, independently of two offers, in effect delegating the decision to her. In particular, Theorem 3 implies that if expert  $i$  chooses constant markup  $b_i$  and the other expert chooses constant markup  $-b_i$  then the principal is always indifferent between the two offers, and she might as well always choose expert  $i$ . Given this strategy of the principal, expert  $i$ 's best response is choosing exactly markup  $b_i$ , which in expectation implements his ideal action. The other expert has no profitable deviation since his proposal is never accepted. While such simple delegation equilibrium exists for each of the two experts, in case of differently biased experts it is more natural to consider the one in which the principal always chooses the expert with the smaller absolute bias, who is expert 2 by convention. These observations are summarized in the next proposition.

**Proposition 1.** *For  $i \in \{1, 2\}$ , an equilibrium exists in which the principal always chooses expert  $i$  and markups are  $k_i = b_i$  and  $k_j = -b_i$  for  $j \neq i$ . The principal's expected payoff in this equilibrium is  $\mathbb{E}[-(s + b_i)^2] = -\sigma^2 - b_i^2$ , where  $s \sim N(0, \sigma^2)$ .*

In the rest of the section we examine equilibria in which the principal's choice between the experts depends in a nontrivial way on the pair of offers proposed.



### 4.3 Upward equilibrium

Here we investigate strategy profiles  $\{(k_1, k_2, a \in \arg \min\{a_1, a_2\}) : k_1 + k_2 \geq 0\}$ . We start by computing players' payoffs under such strategy profiles. Let  $b(k_1, k_2, L) = \mathbb{E}(a - \theta)$  denote the expected bias of the chosen offer and  $Var(k_1, k_2, L) = \text{Var}(a - \theta)$  denote the variance of the chosen offer. Then the principal's utility is:  $V(k_1, k_2, L) = -\mathbb{E}(a - \theta)^2 = -b^2(k_1, k_2, L) - Var(k_1, k_2, L)$ . Proposition 2 below provides the expanded forms of these expressions, which are useful to our analysis. Here and throughout the rest of the paper, let  $f$  and  $F$  denote the PDF and the CDF of the distribution  $N(0, 2\sigma^2)$  and let  $z = k_1 - k_2$ . Note that the expected bias  $b(k_1, k_2, L)$  is strictly less than the expected value of the selected markup,  $k_1(1 - F(z)) + k_2F(z)$ ; this is because the lower offer is associated with a noise term which is normally distributed but truncated above and thus has negative expectation.

**Proposition 2.** *If both experts follow constant markup strategies  $a_j(s_j) = s_j + k_j$  and the principal always chooses the lower offer, then*

$$\begin{aligned} b(k_1, k_2, L) &= -2\sigma^2 f(z) + k_1(1 - F(z)) + k_2F(z); \\ Var(k_1, k_2, L) &= \sigma^2 - 4\sigma^4 f^2(z) - 2\sigma^2 z f(z)(2F(z) - 1) + z^2 F(z)(1 - F(z)); \\ V(k_1, k_2, L) &= -\sigma^2 + 2(k_1 + k_2)\sigma^2 f(z) - k_1^2(1 - F(z)) - k_2^2 F(z); \\ U_i(k_1, k_2, L) &= -\sigma^2 + 2\sigma^2(k_i + k_j - 2b_i)f(z) - (k_j - b_i)^2 F(k_i - k_j) + (B - (k_i - b_i)^2)F(k_j - k_i). \end{aligned}$$

Denote the hazard rate  $\frac{f(x)}{1-F(x)}$  by  $v(x)$  and let  $w(x) := \frac{f(x)}{F(x)}$ . The hazard rate plays an important role in our analysis. It represents the instantaneous probability that experts' signals differ by  $x$ , conditional on differing by at least  $x$ . The following lemma provides several useful bounds. They are the immediate corollaries of [Sampford \(1953\)](#).

**Lemma 6.** *The following inequalities hold for all  $x \in \mathbb{R}$ :*

- $0 < v'(x) < \frac{1}{2\sigma^2}$ ;
- $0 > w'(x) > -\frac{1}{2\sigma^2}$ ;
- $v''(x) > 0$ .

In upward equilibrium, experts are subject to two driving forces acting in opposite directions. The principal's choice of the lower offer implies that offers affected by negative realizations of noise are accepted more frequently. Therefore conditional on being chosen, an expert must revise his belief upwards. This induces experts to increase their markups, similarly to bidders shading their bids downwards in a common value auction environment.<sup>20</sup> At the same time, the bonus (in case  $B > 0$ ) induces experts to decrease their markups in order to be selected more frequently. Hence, if the bonus is small, the first force prevails, and the expert sets a markup higher than her bias. If the bonus is large then the second force prevails, and the expert sets a markup lower than his bias.

<sup>20</sup>A similar shading behavior emerges in the welfare-maximizing (first best) outcome in [Kartik et al. \(2015\)](#), as interestingly the latter outcome implies always selecting the politician with the larger signal in absolute terms.

These conclusions are confirmed by formulas in Theorem 4. If  $B < 2\sigma^2$ , then the markups are higher than the corresponding biases, and if  $B > 2\sigma^2$ , the situation reverses. When the bonus becomes sufficiently high, markups become lower than the biases to an extent that the sum of markups is negative. Then the principal prefers to choose the higher offer. Therefore, for very high bonuses upward equilibria do not exist.

Let  $z^*$  denote the unique solution<sup>21</sup> to the equation

$$z - \rho[v(z) - w(z)] = b_1 - b_2, \quad (4)$$

where we define  $\rho := \sigma^2 - \frac{B}{2}$ .

**Theorem 4.** *There exists a threshold  $B_U > 0$  such that an upward equilibrium exists if and only if  $B \leq B_U$ . When it exists, it is unique and characterized by markups  $k_1^U = b_1 + \rho v(z^*)$  and  $k_2^U = b_2 + \rho w(z^*)$  with  $k_1^U - k_2^U = z^* \geq 0$ . Moreover,  $B_U$  lies in the interval  $[2\sigma^2 + 2\sqrt{\pi}\sigma \max(0, b_2), 2\sigma^2 + \sqrt{\pi}\sigma(b_1 + b_2)]$ . For  $B \leq B_U$ ,*

- $z^* \geq b_1 - b_2 \iff B \leq 2\sigma^2 \iff \rho \geq 0$ ;
- $b(k_1^U, k_2^U, L) = b_U := b_1(1 - F(z^*)) + b_2F(z^*) - Bf(z^*)$ ;
- $\text{Var}(k_1^U, k_2^U, L) = \sigma^2 - 4\sigma^4 f^2(z^*) - 2\sigma^2 z^* f(z^*)(2F(z^*) - 1) + (z^*)^2 F(z^*)(1 - F(z^*))$ .

Notice that the equilibrium markup difference  $z^*$  as well as  $\text{Var}(a - \theta)$  depend on the biases of the experts only through  $b_1 - b_2$ . In upward equilibrium expert 1 wins the bonus with probability  $1 - F(z^*)$ , which is reflected in the expression for  $b_U$ . The bonus affects the expected bias in two ways: through the change in experts' probabilities of winning and through the change in markups. In the appendix (Corollary A.1), we give the full expansion of the players' utilities in upward equilibrium.

In the case of equally biased experts, we can obtain a closed form solution for strategies in upward equilibrium.

**Proposition 3.** *Consider  $b_1 = b_2 = b > 0$ . Then  $B_U = 2\sigma^2 + 2\sqrt{\pi}\sigma b$ , and in upward equilibrium, we have the following:*

- $k_1^U = k_2^U = k_U = b + \frac{\rho}{\sigma\sqrt{\pi}}$ ;
- $b(k_U, k_U, L) = b - \frac{B}{2\sqrt{\pi}\sigma}$ ;
- $\text{Var}(k_U, k_U, L) = \left(1 - \frac{1}{\pi}\right) \sigma^2$ ;
- $V(k_U, k_U, L) = -\left(b - \frac{B}{2\sqrt{\pi}\sigma}\right)^2 - \sigma^2 + \frac{\sigma^2}{\pi}$ ;
- $U_i(k_U, k_U, L) = -\rho + \frac{\sigma^2}{\pi} - \frac{B^2}{4\pi\sigma^2}$  for  $i = 1, 2$ .

---

<sup>21</sup>That there is a unique solution is shown in the proof of Theorem 4.

## 4.4 Downward equilibrium

Recall that an equilibrium is a downward equilibrium when it belongs to  $\{(k_1, k_2, a \in \arg \max(a_1, a_2)) : k_1 + k_2 \leq 0\}$ . Proposition 4 is analogous to Proposition 2. Note that  $b(k_1, k_2, H)$  now exceeds the expected chosen markup because it is associated with a normally distributed noise term truncated from below.

**Proposition 4.** *If both experts follow constant markup strategies  $a_j(s_j) = s_j + k_j$  and the principal always chooses the higher offer, then*

$$\begin{aligned} b(k_1, k_2, H) &= k_1 + k_2 - b(k_1, k_2, L); \\ \text{Var}(k_1, k_2, H) &= \text{Var}(k_1, k_2, L); \\ V(k_1, k_2, H) &= V(-k_1, -k_2, L); \\ U_i(k_1, k_2, H) &= -\sigma^2 - 2\sigma^2(k_i + k_j - 2b_i)f(z) + (B - (k_i - b_i)^2)F(k_i - k_j) - (k_j - b_i)^2F(k_j - k_i). \end{aligned}$$

In downward equilibrium there are again two forces driving experts' choices, similarly to the upward equilibrium scenario, but the directions are reversed. The principal's strategy of choosing the higher offer implies that, conditional on being chosen, an expert must revise his belief and his markup downward. On the other hand, experts want to raise their markups to increase their chances of getting the bonus.

Downward equilibrium differs from upward equilibrium in the following way. Since the sum of biases is positive, the first force must sufficiently outweigh the second force to ensure that the sum of markups is negative, so that the principal's choice of the higher offer is a best response. Hence, noise must be sufficiently large relative to the bonus for downward equilibrium to exist. Even if the bonus is zero, downward equilibrium does not exist for certain parameter values.

The next theorem characterizes downward equilibrium. Note that if  $B_D < 0$ , no downward equilibrium exists. Recall the definition of  $z^*$  from (4).

**Theorem 5.** *There exists a threshold  $B_D$  such that a downward equilibrium exists if and only if  $B \leq B_D$ . When it exists, it is unique and characterized by  $k_1^D = b_1 - \rho w(z^*)$  and  $k_2^D = b_2 - \rho v(z^*)$ , with  $k_1^D - k_2^D = z^* \geq b_1 - b_2$  and  $B_D \in [2\sigma^2 - \sqrt{\pi}\sigma(b_1 + b_2), 2\sigma^2 - 2\sqrt{\pi}\sigma \max(0, b_2)]$ . For  $B \leq B_D$ ,*

- $b(k_1^D, k_2^D, H) = b_D = b_1 F(z^*) + b_2(1 - F(z^*)) + Bf(z^*);$
- $\text{Var}(k_1^D, k_2^D, H) = \sigma^2 - 4\sigma^4 f^2(z^*) - 2\sigma^2 z^* f(z^*)(2F(z^*) - 1) + (z^*)^2 F(z^*)(1 - F(z^*)).$

Note that the equilibrium markup difference  $z^*$  and  $\text{Var}(a - \theta)$  stay the same as in upward equilibrium. In downward equilibrium expert 1 wins the bonus with probability  $F(z^*)$ , which is higher than  $1 - F(z^*)$  in upward equilibrium, and this shifts the expected bias higher. Furthermore, in downward equilibrium the bonus motivates experts to increase their markups, as opposed to upward equilibrium in which the bonus motivates experts to decrease their markups. As a result of the above effects, the expected bias is higher in downward equilibrium. In Corollary A.2 we provide expressions for the players' utilities.

The following corollaries are immediate from Theorems 4 and 5.

**Corollary 1.** *A downward equilibrium exists only if an upward equilibrium exists; that is,  $B_D \leq B_U$ .*

**Corollary 2.** *For  $B \leq B_D$ ,  $k_1^U - b_1 = b_2 - k_2^D$  and  $k_2^U - b_2 = b_1 - k_1^D$ .*

As before, we can obtain a closed form solution for the case of equally biased experts.

**Proposition 5.** *Consider  $b_1 = b_2 = b > 0$ . Then  $B_D = 2\sigma^2 - 2\sqrt{\pi}\sigma b$ , and in downward equilibrium, we have the following:*

- $k_1^D = k_2^D = k_D = b - \frac{\rho}{\sigma\sqrt{\pi}}$ ;
- $b(k_D, k_D, H) = b + \frac{B}{2\sqrt{\pi}\sigma}$ ;
- $Var(k_D, k_D, H) = \left(1 - \frac{1}{\pi}\right) \sigma^2$ ;
- $V(k_D, k_D, H) = -\left(b + \frac{B}{2\sqrt{\pi}\sigma}\right)^2 - \sigma^2 + \frac{\sigma^2}{\pi}$ ;
- $U_i(k_D, k_D, H) = -\rho + \frac{\sigma^2}{\pi} - \frac{B^2}{4\pi\sigma^2}$  for  $i = 1, 2$ .

## 5 Welfare Analysis

In this section we compare the principal's expected utility in upward and downward equilibria, and in the case of simple delegation. We also investigate how the principal's expected payoff in equilibrium depends on the biases of the experts, as well as on the amount of the bonus payment and on the precision of the experts' signals.

### 5.1 Results for non-identical experts

First we establish that whenever both upward and downward equilibria exist, the principal always prefers the former.

**Proposition 6.** *For any fixed  $B \geq 0$ , the principal prefers upward equilibrium to downward equilibrium whenever both exist.*

The intuition for the above result can be summarized as follows. As we pointed out earlier, in any state  $\theta$  variances of the expected offer  $\text{Var}(a-\theta)$  in upward and downward equilibria coincide, but the expected bias  $\mathbb{E}(a-\theta)$  is higher in downward equilibrium:  $b_D = b_1 F(z^*) + b_2(1 - F(z^*)) + Bf(z^*) \geq b_1(1 - F(z^*)) + b_2 F(z^*) - Bf(z^*) = b_U$ . Hence, to conclude that the principal is better off in upward equilibrium it is enough to show that  $|b_D| \geq |b_U|$  or, taking into account above,  $b_U + b_D \geq 0$ . Markup differences in upward and downward equilibria coincide, but the choice rule is opposite, therefore for both experts the probabilities of winning in upward and downward equilibria are complementary. The direct effects of the bonus on  $b_U$  and  $b_D$  are opposite and equal in absolute value. Therefore  $b_U + b_D = b_1 + b_2 \geq 0$ , implying that the principal is better off in upward equilibrium.

In the Supplementary Appendix we show that in case the two experts are not equally biased, the expert with the lower bias also prefers upward equilibrium to downward equilibrium, while the expert with the higher bias has the opposite preferences. This is both because the expected action is closer

to expert 2's ideal point in upward equilibrium, and closer to expert 1's ideal point in downward equilibrium, and because expert 2 is chosen (and hence receives the bonus) with higher probability in upward equilibrium, and expert 1 is chosen with higher probability in downward equilibrium.

The above result motivates us to focus on comparing the principal's utility in upward equilibrium versus in simple delegation to the expert with the smaller absolute bias. In general this comparison is complicated, but in the next proposition we show that for upward equilibrium to be better for the principal, it has to be the case that either both experts have high enough biases, or that they are equally biased.

**Proposition 7.** *Consider  $B = 0$  and biases  $b_1 = b + x$  and  $b_2 = b - x$  for some  $x, b \geq 0$ . If  $x > 0$ , there exists a threshold  $\bar{b} > 0$  such that the principal prefers simple delegation to the upward equilibrium if and only if  $b_1 > \bar{b}$ . If  $x = 0$ , the principal always prefers the upward equilibrium.*

*Proof.* Delegation to expert 2 alone yields principal utility  $V^S(b, x) = -\sigma^2 - (b - x)^2$ , while upward equilibrium yields

$$V^U(b, x) = -\sigma^2 - b^2 - x^2 + (2F(z^*) - 1)2bx + (k_1 - b - x)(k_2 - b + x).$$

The difference between these is

$$\begin{aligned} V^U(b, x) - V^S(b, x) &= -2(1 - F(z^*))bx + (k_1 - b - x)(k_2 - b + x) \\ &= -2(1 - F(z^*))bx + \sigma^4 v(z^*)w(z^*). \end{aligned}$$

Recall that  $z^*$  is independent of  $b$ . If  $x = 0$ , then the above expression is always positive. For fixed  $x > 0$  then the existence of  $\bar{b}$  follows from the fact that this expression is decreasing linearly in  $b$  and positive for  $b = 0$ .  $\square$

For intuition behind the above result, first consider the case  $x > 0$ . The variance of the chosen action in upward equilibrium is always lower than that in simple delegation, and both are independent of  $b$ . As  $b$  increases, the expected bias  $b - (2F(z^*) - 1)x$  in upward equilibrium and  $b - x$  in downward equilibrium increase at the same rate. Since the former is higher than the latter, and losses are quadratic, this increase hurts more in upward equilibrium than in simple delegation. Once  $b$  is high enough, this disadvantage outweighs the initial advantage of lower variance.

For equal biases, the intuition for why upward equilibrium, when it exists, yields a higher expected payoff to the principal than simple delegation generalizes for  $B \geq 0$  as follows. With simple delegation, the expected utility of the principal is  $V = -b^2 - \sigma^2$ . In contrast, in upward equilibrium the expected bias of the action is  $b - \frac{B}{2\sqrt{\pi}\sigma} < b$ , and the variance is  $\sigma^2 - \frac{\sigma^2}{\pi} < \sigma^2$ . The decreased variance of the action implies that the principal prefers upward equilibrium to simple delegation, even when  $B = 0$  and therefore the expected action is the same in upward equilibrium as in simple delegation.

We conclude this subsection by comparing the expected payoff the principal can achieve with two equally biased experts,  $b_1 = b_2 = b$ , to the expected payoff she can achieve with two oppositely biased experts,  $b_1 = -b_2 = b$ . Theorem 4 implies that the expected bias of the action in the

case of equally biased experts is equal to  $b - Bf(0)$ , simplifying to  $b$  when  $B = 0$ . In the case of oppositely biased experts the expected bias of the action is  $b(1 - 2F(z^*)) - Bf(z^*)$ , simplifying to  $b(1 - 2F(z^*))$  when  $B = 0$ . Hence, with  $B = 0$  the absolute value of the expected bias in the case of oppositely biased experts is lower than in the case of equally biased experts. However, the next Proposition shows that the variance of the action is lower in the symmetric case, and in fact this effect dominates, resulting in the principal preferring to have two equally biased experts. Let  $V_{symm}(b)$  be the principal's expected payoff in upward equilibrium when  $b_1 = b_2 = b$ , let  $V_{opp}(b)$  be the principal's expected payoff in upward equilibrium when  $b_1 = -b_2 = b$ , and let  $V_{sim}(b) = -\sigma^2 - b^2$  be the principal's expected payoff in case of simple delegation to an expert with absolute bias  $b$ .

**Proposition 8.** *For any  $b > 0$ ,  $V_{symm}(b) \geq V_{opp}(b) \geq V_{sim}(b)$ .*

*Proof.* Upward equilibrium for opposite biases exists only if  $B \leq B_U = 2\sigma^2$ . From Theorem 4

$$\begin{aligned} V_{symm}(b) &= -\sigma^2 - b^2 + 2Bbf(0) + (4\sigma^4 - B^2)f^2(0) \\ V_{opp}(b) &= -\sigma^2 - b^2 + \left(\sigma^4 - \frac{B^2}{4}\right) \frac{f(z_{opp}^*)}{F(z_{opp}^*)(1 - F(z_{opp}^*))} \geq -\sigma^2 - b^2 = V_{sim}(b), \end{aligned}$$

where  $z_{opp}^*$  is the upward equilibrium markup difference in the case of oppositely biased experts. As  $2Bbf(0) \geq 0$  and  $\frac{f^2(z)}{F(z)(1-F(z))}$  reaches its maximum at  $z = 0$ , we get  $V_{symm}(b) \geq V_{opp}(b)$   $\square$

## 5.2 Equally Biased Experts

In this subsection we focus on the case when the experts are equally biased:  $b_1 = b_2 = b > 0$ . In this setting we can derive closed form solutions for the marginal affects of various parameter values on the principal's expected payoff, greatly simplifying comparative statics exercises.

Recall from Proposition 7 that in the case of equal biases, the principal always prefers upward equilibrium to simple delegation to one of the biased experts. Moreover, using the formula derived in Proposition 8 for the principal's expected utility in upward equilibrium, we can exactly characterize when it is the case that the principal prefers upward equilibrium with two equally biased experts to simple delegation to an unbiased expert.

**Corollary 3.** *The principal prefers upward equilibrium with equally biased experts to unconstrained delegation to an unbiased expert if and only if  $B \in [2\sqrt{\pi}\sigma b - 2\sigma^2, B_u]$ . For  $B = 0$  this condition is equivalent to bias-to-noise ratio being low enough:*

$$\frac{b}{\sigma} \leq \frac{1}{\sqrt{\pi}}.$$

*Proof.* It is enough to compare  $V_{symm} = -(b - \frac{B}{2\sqrt{\pi}\sigma})^2 - \sigma^2 + \frac{\sigma^2}{\pi}$  with utility from single delegation to unbiased expert  $-\sigma^2$ .  $\square$

From here on we investigate comparative statics of the principal's expected payoff in upward equilibrium. First we look at how the principal's expected payoff depends on  $b$ , the common bias of the experts. Recall that upward equilibrium exists only for  $B \leq B_u = 2\sigma^2 + 2\sqrt{\pi}\sigma b$ . The principal's

expected payoff in upward equilibrium is  $V(k_U, k_U, L) = -(b - \frac{B}{2\sqrt{\pi}\sigma})^2 - (1 - \frac{1}{\pi})\sigma^2$ . This expression is maximized at  $b^* = \frac{B}{2\sqrt{\pi}\sigma}$ , taking the value  $-(1 - \frac{1}{\pi})\sigma^2$ . This implies that for  $B = 0$ , it is optimal for the principal to have nonbiased experts, but for  $B > 0$ , the optimal bias level is strictly positive and increasing in  $B$ .

Next we consider how the principal's expected payoff depends on the precision of the experts' signals. The expected bias of the action  $b_U = b - \frac{B}{2\sqrt{\pi}\sigma}$  is increasing in  $\sigma$ . Nevertheless, the equilibrium utility of the principal,  $V(k_U, k_U, L) = -(b - \frac{B}{2\sqrt{\pi}\sigma})^2 - \sigma^2 + \frac{\sigma^2}{\pi}$ , is nonmonotonic in  $\sigma$ , for the following reason. If  $\sigma$  increases, then the expected variance of the action  $(1 - \frac{1}{\pi})\sigma^2$  increases, leading to a decrease in the principal's expected payoff. At the same time when the variance of noise  $\sigma$  is small, the value of the expected bias  $b - \frac{B}{2\sqrt{\pi}\sigma}$  stays negative and increases, leading to a decrease in absolute value of the expected bias, which increases the principal's expected payoff. After  $\sigma$  reaches the value when  $B$  is optimal in exogenous case ( $B = 2\sqrt{\pi}\sigma b$ ) the value of expected bias turns positive and further increase in  $\sigma$  leads to decrease in utility. Hence, optimal value for principal lies on the interval  $(0, \frac{B}{2\sqrt{\pi}b})$ . The next proposition provides the formal comparative statics of the principal's utility in the precision of the experts' signals.

Let  $B > 0$  and  $\sigma^*$  denote the unique positive solution to the equation  $(4\pi - 4)\sigma^4 + 2\sqrt{\pi}bB\sigma = B^2$ .

**Proposition 9.** *Consider  $b_1 = b_2 = b > 0$ .*

*If  $0 < B < \frac{2(\pi-1)\pi}{(\pi-2)^2}b^2$ , there exists a non-empty interval  $\sigma \in [\frac{\sqrt{\pi b^2 + 2B} - \sqrt{\pi}b}{2}, \sigma^*)$ , where upward equilibrium exists and the principal's expected payoff is increasing in  $\sigma$ ; when  $\sigma > \sigma^*$ , the principal's expected payoff is decreasing in  $\sigma$ .*

*If  $B = 0$  or  $B \geq \frac{2(\pi-1)\pi}{(\pi-2)^2}b^2$ , the principal's expected payoff is always decreasing in  $\sigma$ .*

For a concrete example when the principal's expected payoff increases in  $\sigma$ , consider  $b_1 = b_2 = b = 2$  and  $B = 9 < \frac{8(\pi-1)\pi}{(\pi-2)^2}$ . Then  $V(\sigma)$  is increasing on an interval containing  $(0.992, 1.085)$ .

Lastly, we address the question of how the principal's expected payoff depends on the bonus payment. We start with the case when the bonus payment comes from exogenous sources, and therefore only indirectly affecting the principal's payoff, through influencing experts' strategies in upward equilibrium. While the variance of the action  $Var(k_U, k_U, L) = \sigma^2 - \frac{\sigma^2}{\pi}$  does not depend on  $B$ , the expected bias  $b_U = b - \frac{B}{2\sqrt{\pi}\sigma}$  is decreasing in  $B$ . The principal prefers the expected bias to be as close to 0 as possible, so her expected payoff in upward equilibrium is maximized at  $B = 2\sqrt{\pi}\sigma b$ , where it is equal to  $-\sigma^2[1 - \frac{1}{\pi}]$ . At  $B = 0$ ,  $b_U = b - \frac{B}{2\sqrt{\pi}\sigma} = b \geq 0$  and a small increase in bonus decreases the experts' markup and benefits the principal. However, at  $B = B_u$   $b_U = b - \frac{B}{2\sqrt{\pi}\sigma} = -\frac{\sigma}{\sqrt{\pi}} < 0$  and principal prefers to increase markups and correspondingly lower bonus. As a consequence, an intermediate point  $B = 2\sqrt{\pi}\sigma b$  is optimal.

Next we consider the situation, when the bonus is paid from the principal's pocket. For this investigation, we append our game (described in Section 2) with a stage 0, preceding stage 1, in which the principal chooses  $B$ . We assume that the principal's choice of  $B$  becomes public knowledge by stage 1. We also modify the principal's payoff to  $V(a, \theta) = -(a - \theta)^2 - B$ , corresponding to the assumption that the principal has to pay the bonus.

The next proposition characterizes the optimal bonus choice of the principal in this case.



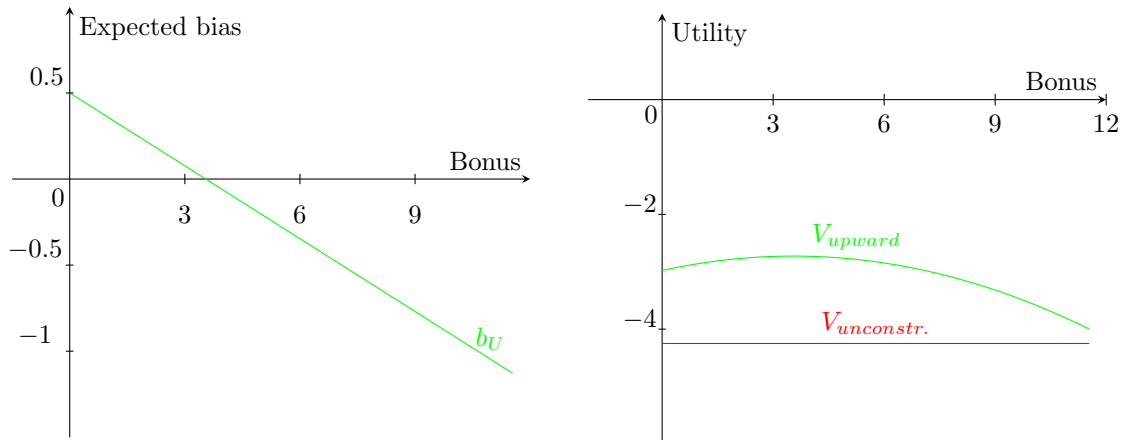


Figure 2: The Expected Bias and the Principal's Payoff (exogenous bonus) for  $b = 0.5, \sigma = 2$

**Proposition 10.** *Suppose that bonus  $B$  is paid by the principal. The principal's optimal choice of  $B$  depends on the bias-to-noise ratio.*

*If  $\frac{b}{\sigma} \leq \sqrt{\pi}$ , then the principal pays no bonus:  $B = 0$ .*

*If  $\frac{b}{\sigma} > \sqrt{\pi}$ , then the principal chooses bonus  $B = 2\sqrt{\pi}\sigma^2 \left[\frac{b}{\sigma} - \sqrt{\pi}\right]$ . The increase in the principal's expected payoff, relative to when the bonus is restricted to be 0, is equal to  $(b - \sqrt{\pi}\sigma)^2$ .*

Therefore, if the bias-to-noise ratio is high enough, it is optimal for the principal to offer a positive bonus.

At  $B = B_u$ :  $b_U = k_U - \frac{\sigma}{\sqrt{\pi}} = -\frac{\sigma}{\sqrt{\pi}} < 0$  and the principal would like to decrease the bonus, in order to increase experts' markups and drive the expected bias closer to 0.

At  $B = 0$ :  $b_U = b \geq 0$ , and if the principal increases the bonus then the expected bias decreases. Therefore a marginal increase of  $B$  from 0 improves the principal's expected payoff whenever the marginal gain from the decrease in expected bias ( $\frac{b}{\sqrt{\pi}\sigma}$ ) exceeds the marginal expense from increasing the bonus (1).

Proposition 10 is illustrated on Figure 3: the principal chooses a strictly positive bonus if the decrease in expected bias exceeds the marginal disutility from increasing bonus:  $\frac{b}{\sqrt{\pi}\sigma} > 1$ . This for example holds for  $b = 2, \sigma = 0.5$ , but not for  $b = 1, \sigma = 1$ .

## 6 Extensions

### 6.1 Ex Ante Commitment by the Principal

The analysis in the previous sections assumes that the principal plays a best response to the strategies of the experts. Alternatively, we can consider situations in which the principal can ex ante commit to a simple strategy. In particular, assume that the principal can credibly commit to choose the lower offer with any probability  $p \in [0, 1]$  and the higher offer with probability  $1 - p$ . We restrict attention

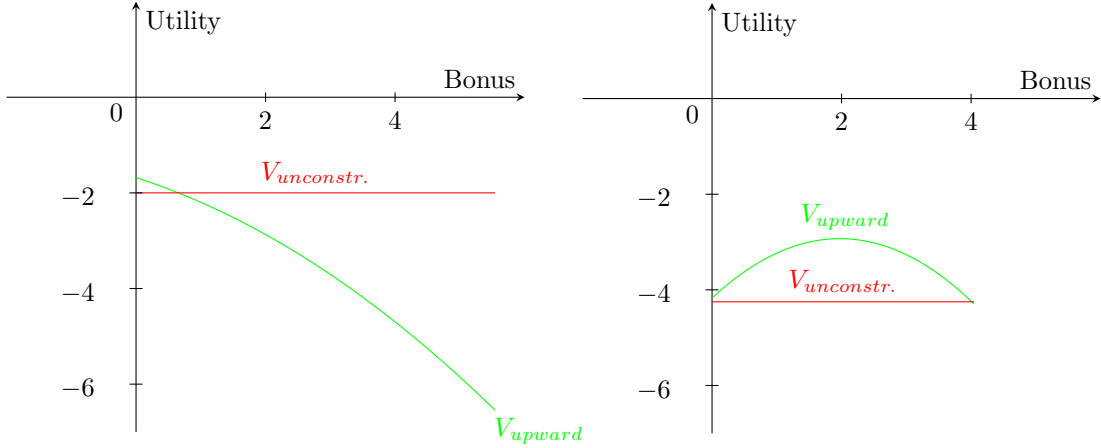


Figure 3: The Principal's Payoff (bonus paid by Principal) for  $b = 1$ ,  $\sigma = 1$  on the left; for  $b = 2$ ,  $\sigma = 0.5$  on the right

to the case of equally biased experts:  $b_1 = b_2 = b$ . While for very large bonuses such a commitment opportunity might be beneficial for the principal, here we show that for  $B \leq 2\sigma^2$  commitment does not improve the principal's payoff, relative to the expected payoff in upward equilibrium in the game without commitment.

If the principal commits to choose the lower offer with probability  $p$ , the experts' FOCs are

$$k_1 = b + \rho \frac{f(z)(2p-1)}{p(1-F(z)) + (1-p)F(z)}$$

$$k_2 = b + \rho \frac{f(z)(2p-1)}{pF(z) + (1-p)(1-F(z))},$$

where  $z = k_1 - k_2$ .

As  $B \leq 2\sigma^2$ , by Lemma S.1 from Supplementary Appendix, the only equilibrium of the game between the two experts, given the pre-committed strategy of the principal, involves  $k_1 = k_2 = k = b + (2p-1)\frac{2\sigma^2-B}{2\sqrt{\pi}\sigma}$ . The principal's utility is

$$V = p \left[ -\sigma^2 - k^2 + \frac{2\sigma}{\sqrt{\pi}}k \right] + (1-p) \left[ -\sigma^2 - k^2 - \frac{2\sigma}{\sqrt{\pi}}k \right]$$

$$= \frac{\rho(2\sigma^2 + B)}{2\pi\sigma^2} (2p-1)^2 + \frac{bB}{\sqrt{\pi}\sigma} (2p-1) - \sigma^2 - b^2.$$

The next proposition follows from inspection of the expression above.

**Proposition 11.** *For  $B \leq 2\sigma^2$ , the optimal strategy of the principal under commitment is  $p = 1$ .*

In particular, commitment by the principal results in the same outcome as in the upward equilibrium of the game without commitment.

## 6.2 Unselected Expert Indifferent over Actions

In the baseline model we assumed that an expert whose offer is not selected is still affected by principal's action. While this is a reasonable assumption in some contexts, in other situations it is more realistic to assume that the expert not selected by the principal receives an outside payoff that is independent of the state and the implemented action. For instance, a car mechanic is unlikely to care what kind of maintenance is done if he is not the one selected for the job. In this extension we assume that if expert  $i$  is chosen then expert  $j$ 's realized payoff is normalized to be 0. We restrict attention the case of equally biased experts:  $b_1 = b_2 = b > 0$

In this version of the model we assume that  $B$  is large enough that an expert's expected payoff under simple delegation to that expert is nonnegative; that is, he prefers simple delegation to not being selected at all. Under this condition, the same simple delegation equilibria exist in this version of the model as in the baseline model. Below we show that under some conditions there also exist symmetric pure strategy equilibria that are similar to the ones characterized in the baseline model. For this to be the case, the bonus payment must be neither too low nor too large.

First we examine the conditions for the existence of upward equilibrium. Using the same notation as before, we investigate strategy profiles  $\{(k_1, k_2, C(a_1, a_2) \in \arg \min\{a_1, a_2\}) : k_1 + k_2 \geq 0\}$ . While for any such profile the principal's payoff does not change, the experts' expected payoffs should be recalculated:

$$\begin{aligned} U_i(k_i, k_j, L) &= \int_{k_i - k_j}^{\infty} \left[ B - \left( k_i - b - \frac{t}{2} \right)^2 - \frac{\sigma^2}{2} \right] f(t) dt \\ &= [B - \sigma^2 - (k_i - b)^2] (1 - F(k_i - k_j)) + \left[ 2\sigma^2(k_i - b) - \frac{1}{2}\sigma^2(k_i - k_j) \right] f(k_i - k_j). \end{aligned}$$

Notice that for a fixed constant markup strategy of the other expert, an expert can choose arbitrarily high constant markup and guarantee an expected payoff arbitrarily close to 0.<sup>22</sup> Hence, 0 is a lower bound for experts' equilibrium payoffs.

In what follows, define  $\beta := \sqrt{\pi + \frac{B}{\sigma^2} - \frac{5}{2}}$ .

**Proposition 12.** *A symmetric upward equilibrium  $k_1^U = k_2^U = k_U$  exists if and only if  $B \in \left[ \left( \frac{5}{2} - \frac{3\pi(8\pi-11)}{16(\pi-1)^2} \right) \sigma^2, \frac{5}{2}\sigma^2 + 2\sqrt{\pi}b\sigma + b^2 \right]$ . When it exists, it is characterized by:*

- $k_1^U = k_2^U = k_U = b + (\sqrt{\pi} - \beta) \sigma$ ;
- $b(k_U, k_U, L) = b + \left( \sqrt{\pi} - \beta - \frac{1}{\sqrt{\pi}} \right) \sigma$ ;
- $Var(k_U, k_U, L) = \left( 1 - \frac{1}{\pi} \right) \sigma^2$ ;
- $V(k_U, k_U, L) = - \left[ b + \left( \sqrt{\pi} - \beta - \frac{1}{\sqrt{\pi}} \right) \sigma \right]^2 - \sigma^2 + \frac{\sigma^2}{\pi}$ ;
- $U_i(k_U, k_U, L) = \left[ \frac{\pi-1}{\sqrt{\pi}} \beta + \frac{7}{4} - \pi \right] \sigma^2$  for  $i = 1, 2$ .

<sup>22</sup>For this reason, we do not introduce an explicit participation constraint in this version of the model, even though such a constraint would be natural in many applications.

In the appendix we show that  $k_U = b + (\sqrt{\pi} - \beta)\sigma > k_U^{bas.} = b + (1 - \frac{B}{2\sigma^2})\frac{\sigma}{\sqrt{\pi}}$ , hence in this version of the model experts select higher markups in upward equilibrium than in the baseline model (for parameter values for which upward equilibrium exists in both model versions). The intuition behind this result is that in this alternative version of the model, the relative gain from being selected is reduced by the policy loss (that is not imposed on the expert if not selected). Since we consider  $B$  sufficiently large that expected payoffs are nonnegative, the resulting “net bonus” is still nonnegative; being selected is still preferable, conditional on having made the lower offer. It follows that an expert’s equilibrium offer in either version is lower than what is ex-post optimal for that expert – that is, optimal after conditioning on both the expert’s signal and having the lower offer. The smaller net bonus in the alternative version reduces the expert’s incentive to marginally lower his offer in order to more frequently earn the net bonus. This reduction must be met by an offsetting reduction in his incentive to raise his offer, which is enforced through his bidding higher and thus closer to his ex-post optimum; due to quadratic losses, marginal movements toward the ex-post optimum have decreasing marginal benefits.

We note that the qualitative comparison between this extension and the baseline model is dependent upon the modeling of preferences over policy outcomes through losses. Such a model is appropriate for applications where an expert would prefer not to be associated with the project if his action would be sufficiently far from the true state; for example, this would be the case if the expert has a reputation at stake. Alternatively, one could model preferences through gains, using some single-peaked, nonnegative utility function of the distance between the action and the true state. In that model, the comparison above would be reversed, as being selected enhances the bonus and thus experts compete more aggressively by lowering their offers.

Next we turn attention to characterizing the conditions under which a downward equilibrium exists in which the principal always chooses the higher offer. Experts’ expected payoffs can be calculated as:

$$\begin{aligned} U_i(k_i, k_j, H) &= \int_{-\infty}^{k_i - k_j} \left[ B - \left( k_i - b - \frac{t}{2} \right)^2 - \frac{\sigma^2}{2} \right] f(t) dt \\ &= [B - \sigma^2 - (k_i - b)^2] F(k_i - k_j) - \left[ 2\sigma^2(k_i - b) - \frac{1}{2}\sigma^2(k_i - k_j) \right] f(k_i - k_j) \end{aligned}$$

As in upward equilibrium, 0 is a lower bound for experts’ equilibrium payoffs.

**Proposition 13.** *Consider  $b_1 = b_2 = b > 0$ . If  $\frac{b}{\sigma} > \frac{3\sqrt{\pi}}{4(\pi-1)}$ , then no symmetric downward equilibrium exists. If  $\frac{b}{\sigma} \leq \frac{3\sqrt{\pi}}{4(\pi-1)}$ , then a symmetric downward equilibrium  $k_1^D = k_2^D = k_D$  exists if and only if  $B \in \left[ \left( \frac{5}{2} - \frac{3\pi(8\pi-11)}{16(\pi-1)^2} \right) \sigma^2, \frac{5}{2}\sigma^2 + 2\sqrt{\pi}b\sigma + b^2 \right]$ . When it exists, it is characterized by:*

- $k_1^D = k_2^D = k_D = b - (\sqrt{\pi} - \beta)\sigma$ ;
- $b(k_D, k_D, H) = b - \left( \sqrt{\pi} - \beta - \frac{1}{\sqrt{\pi}} \right) \sigma$ ;
- $Var(k_D, k_D, H) = \left( 1 - \frac{1}{\pi} \right) \sigma^2$ ;
- $V(k_D, k_D, H) = - \left[ b - \left( \sqrt{\pi} - \beta - \frac{1}{\sqrt{\pi}} \right) \sigma \right]^2 - \sigma^2 + \frac{\sigma^2}{\pi}$ ;

- $U_i(k_D, k_D, H) = \left[ \frac{\pi-1}{\sqrt{\pi}}\beta + \frac{7}{4} - \pi \right] \sigma^2$  for  $i = 1, 2$ .

For downward equilibrium, the difference relative to the baseline model is the mirror image of the difference described earlier for upward equilibrium. Again, the bonus is reduced by quadratic losses, but in downward equilibrium this causes markups to decrease, as experts compete less aggressively to make the higher offer.

## 7 Conclusion

We proposed a model in which a principal can choose between two imperfectly informed experts, introducing the possibility of competition in a delegation framework. We showed that a principal with limited knowledge of the decision environment can benefit from the presence of two experts, relative to a simple unconstrained delegation to one of them, even if the experts have exactly the same bias. The main reason is that in equilibria in which the selection of the expert depends nontrivially on the experts' proposals, information is utilized from both experts' private signals. The option of offering a bonus payment to the selected expert can improve the principal's payoff, by inducing the experts to report more truthfully, but only to a certain point. The effects of more precise signals to the experts on the principal's payoff is in general ambiguous, but they are always beneficial for small enough bonus payments.

We also find that in the equilibria we characterize, the principal being severely restricted in her choice of strategies, due to limited knowledge of the environment, might not hurt her. In particular, given the equilibrium strategies of the experts, the principal would not be able to do better without the restrictions to simple strategies. Moreover, for small bonuses the best equilibrium of the game achieves the payoff the principal could get by ex ante committing to a strategy. Hence, if there is competition between experts to be selected, sophistication on the part of the experts can lead to the same outcome no matter of the principal's knowledge about the environment.

As this is the first step in investigating the benefits of multiple choices of experts in a delegation problem, there are many avenues of future research. One is examining multi-dimensional environments, in which different experts differ in their dimensions of specialization. Another direction would be investigating the problem of choosing an expert to delegate a task to with a more general mechanism design approach.

## A Appendix

### A.1 Proofs for Section 3.2

*Proof of Lemma 1.* The proof makes use of two facts: (i) for mappings on metric spaces, continuity is equivalent to sequential continuity,<sup>23</sup> and (ii) weak convergence of probability distributions is equivalent to pointwise convergence of the CDFs at all points of continuity.

<sup>23</sup>A mapping  $f : X \rightarrow Y$  is sequentially continuous if for all  $x \in X$ ,  $x_n \rightarrow x$  implies  $f(x_n) \rightarrow f(x)$ . In general, sequential continuity is weaker than continuity.

To show that the mapping  $G_\mu$  is continuous in  $\theta$ , for any  $\theta$ , choose any  $x$  at which  $G_\mu(x; \theta)$  is continuous in  $x$ , and consider any sequence  $\theta_n \rightarrow \theta$ . We claim that  $\mu((-\infty, x - (s_i - \theta)]|s_i)$  is continuous in  $x$ , and hence  $\theta$ , for almost all  $s_i \in S_i$ . Since CDFs are right-continuous, it suffices to show left-continuity. If left-continuity fails for all  $s_i \in D$  for some  $D \subset S_i$  of positive measure, then for all  $s_i \in D$  and all sequences  $(x_n) \nearrow x$ ,  $\mu((-\infty, x_n - (s_i - \theta)]|s_i) \leq \mu((-\infty, x - (s_i - \theta)]|s_i) < \mu((-\infty, x - (s_i - \theta)]|s_i)$ , and thus  $\lim_{n \rightarrow \infty} G_\mu(x_n; \theta) < G_\mu(x; \theta)$ , contradicting continuity of  $G_\mu(x; \theta)$  in  $x$ . Thus for almost all  $s_i \in S_i$ , the integrand in (1) is continuous in  $\theta$ . Thus  $\mu((-\infty, x - (s_i - \theta_n)]|s_i) f(s_i|\theta_n) \rightarrow \mu((-\infty, x - (s_i - \theta)]|s_i) f(s_i|\theta)$  pointwise almost everywhere. Since the  $\mu(\cdot|s_i)$  are probability measures, we have  $\mu((-\infty, x - (s_i - \theta_n)]|s_i) f(s_i|\theta_n) \leq f(s_i|\theta_n)$ . Note that for all  $s_i$ ,  $f(s_i|\theta_n) \rightarrow f(s_i|\theta)$  and  $1 = \int_{-\infty}^{\infty} f(s_i|\theta_n) \lambda(ds_i) \rightarrow \int_{-\infty}^{\infty} f(s_i|\theta) \lambda(ds_i) = 1$ . By application of the generalized Lebesgue Convergence Theorem,<sup>24</sup> we have  $G_\mu(x; \theta_n) = \int_{-\infty}^{\infty} \mu((-\infty, x - (s_i - \theta_n)]|s_i) f(s_i|\theta_n) \lambda(ds_i) \rightarrow \int_{-\infty}^{\infty} \mu((-\infty, x - (s_i - \theta)]|s_i) f(s_i|\theta) \lambda(ds_i) = G_\mu(x; \theta)$ . Hence  $G_\mu(\cdot; \theta_n)$  converges weakly to  $G_\mu(\cdot; \theta)$ , as desired.  $\square$

*Proof of Lemma 2.* We have already noted that these are metric spaces. We first show that  $\mathcal{M}^M$  is a relatively compact. Given any  $\eta > 0$ , choose any compact  $T \subset S_i$  such that  $\gamma(T) > 1 - \eta$ . Since  $X_i^M$  is compact,  $T \times X_i^M$  is compact and  $\mu(T \times X_i^M) = \gamma(T) > 1 - \eta$  for all  $\mu \in \mathcal{M}^M$ . By definition,  $\mathcal{M}^M$  is thus a tight family of measures. By Prokhorov's Theorem,<sup>25</sup>  $\mathcal{M}^M$  is relatively compact. To see that  $\mathcal{M}^M$  is closed, take any sequence  $(\mu_n)_{n \in \mathbb{N}}$  of measures in  $\mathcal{M}^M$  converging weakly to  $\mu \in \mathcal{P}$ ; we show that  $\mu \in \mathcal{M}^M$ . For any closed  $T \subset S_i$ , the set  $T \times X_i^M$  is closed and thus weak convergence implies<sup>26</sup>  $\mu(T \times X_i^M) \geq \limsup_n \mu_n(T \times X_i^M) = \gamma(T)$ . Suppose  $\mu(T \times X_i^M) > \gamma(T)$  for some such  $T$ . If so, then  $\mu(T \times X_i^M) + \mu(T^c \times X_i^M) = 1 = \gamma(T) + \gamma(T^c) \implies \mu(T^c \times X_i^M) < \gamma(T^c) - \eta$  for some  $\eta > 0$ . There exists a closed set  $T' \subset T^c$  such that  $\gamma(T') > \gamma(T^c) - \eta$ , and thus  $\mu(T' \times X_i^M) < \gamma(T')$ , a contradiction. Hence  $\mu(T \times X_i^M) = \gamma(T)$  for all closed  $T \subset S_i$ , which implies that  $\mu \in \mathcal{M}$ . Being both closed and relatively compact,  $\mathcal{M}$  is compact.

That  $\mathcal{G}^M$  is compact follows immediately from the compactness of  $\mathcal{M}^M$  combined with the continuity of  $G : \mu \mapsto G_\mu$  from Lemma 1.  $\square$

*Proof of Lemma 3.* Suppose on the contrary that for each  $y \in Y$ , there exists  $\eta_y > 0$  such that  $\lambda\{x \in \mathbb{R} : d(\pi(x), y) < \eta_y\} < \infty$ . The collection  $\{N_{\eta_y}(y) : y \in Y\}$  is an open covering of  $Y$ , and by compactness, it has a finite subcovering denoted  $\{N_{\eta_i}(y_i)\}_{i=1}^n$  for some  $n \in \mathbb{N}$ . It follows that  $\mathbb{R} \subseteq \cup_{i=1}^n \pi^{-1}(N_{\eta_i}(y_i))$  and thus  $\lambda X \leq \sum_{i=1}^n \lambda \pi^{-1}(N_{\eta_i}(y_i)) < \infty$ , a contradiction.  $\square$

*Proof of Theorem 1.* Given  $\mu \in \mathcal{M}_0$ ,  $\mu \in \mathcal{M}^M$  for some  $M \in \mathbb{N}$ . First, we show that  $\mu$  is nearly stationary, and then we show payoff equivalence. Suppose otherwise that  $\mu$  is not nearly stationary. By Lemma 1,  $G_\mu : \theta \mapsto G_\mu(\cdot; \theta)$  is a continuous, and hence measurable, function from  $\mathbb{R}$  to  $\mathcal{G}^M$ , which is compact by Lemma 2. By Lemma 3, there exists an attraction  $Q \in \mathcal{G}^M$  for  $\mu$ .

Next, given the existence of an attraction  $Q$ , we establish uniqueness. Suppose there also exists an attraction  $Q' \in \mathcal{G}^M$  for  $G_\mu$  with  $Q' \neq Q$ . We show that there exists a (constant markup) strategy

<sup>24</sup>See Royden and Fitzpatrick (1988, Theorem 4.17).

<sup>25</sup>See, for example, Billingsley (2009, Theorem 5.1).

<sup>26</sup>See Billingsley (2009, Theorem 2.1).

of the rival and a simple principal strategy such that if expert 1 plays  $\mu$ , she does not have well-defined payoff in the limit. For concreteness, suppose the following: the principal plays  $C$  defined by  $C(x_1, x_2) = 1$  if  $x_1 \leq x_2$  and 0 otherwise; expert 2 plays a constant markup strategy  $\kappa_m$  for some  $m \in \mathbb{R}$ ; and  $b_1 = 0$ . Let  $\nu_n := G_\mu(\cdot; \theta_n) \otimes G_{\kappa_m}$  and  $\nu := Q \otimes G_{\kappa_m}$ .<sup>27</sup> We show that if  $(\theta_n)$  is a sequence such that  $G_\mu(\cdot; \theta_n) \Rightarrow Q$ , then  $u(\theta_n) \rightarrow u_{Q, \kappa_m} := -\int_{\mathbb{R}^2} (\min\{x_1, x_2\})^2 d\nu(x_1, x_2)$ . To prove the convergence, recall that  $u(\theta_n) = -\int_{\mathbb{R}^2} (\min\{x_1, x_2\})^2 d\nu_n(x_1, x_2)$ . Since  $G_\mu(\cdot; \theta_n) \rightarrow Q$ , we have  $\nu_n \Rightarrow \nu$ . Moreover, we have  $u(\theta_n) \rightarrow u_{Q, \kappa_m} < \infty$ .<sup>28</sup> Likewise, if  $(\theta_n)$  is such that  $G_\mu(\cdot; \theta_n) \Rightarrow Q'$ ,  $u(\theta_n)$  tends to  $u_{Q', \kappa_m}$ , defined analogously to that above. We claim that  $u_{Q, \kappa_m} \neq u_{Q', \kappa_m}$  for some  $m \in \mathbb{R}$ . Suppose not. Note that for any  $Q \in \mathcal{G}_M$ ,  $u_{Q, \kappa_m} = -\int_{\mathbb{R}^2} (\min\{x_1, s_2 + m\})^2 (Q \otimes H)(d(x_1, s_2))$ , where  $H$  is the CDF of  $N(0, \sigma^2)$ . Next, we show that we can differentiate w.r.t.  $m$  under the integral. Write  $\mathbf{x} := (x_1, s_2)$ ,  $g(\mathbf{x}, m) := (\min\{x_1, s_2 + m\})^2$ , and  $\nu = Q \otimes H$ . Note that for all  $\mathbf{x} \in \mathbb{R}^2$ ,  $g(\mathbf{x}, m)$  is absolutely continuous in  $m$  on bounded intervals. We have

$$a(m) := \int_{\mathbb{R}^2} g(\mathbf{x}, m) \nu(d\mathbf{x}) \quad (5)$$

$$= \int_{\mathbb{R}^2} \left[ g(\mathbf{x}, m_0) + \int_{m_0}^m g_m(\mathbf{x}, z) dz \right] \nu(d\mathbf{x}) \quad (6)$$

$$= \int_{\mathbb{R}^2} g(\mathbf{x}, m_0) \nu(d\mathbf{x}) + \int_{m_0}^m \int_{\mathbb{R}^2} g_m(\mathbf{x}, z) \nu(d\mathbf{x}) dz \quad (7)$$

$$\implies a'(m) := \int_{\mathbb{R}^2} g_m(\mathbf{x}, m) \nu(d\mathbf{x}) \quad \text{a.e. } m, \quad (8)$$

where  $m_0 < m$  can be chosen arbitrarily. To obtain (6) we have used the fact that absolutely continuous functions are the integral of their derivatives.<sup>29</sup> Fubini's Theorem is used to obtain (7). Differentiability-a.e. of the integral yields (8).<sup>30</sup> By the same arguments, we obtain  $a''(m) = \int g_{mm}(\mathbf{x}, m) \nu(d\mathbf{x}) = 2 \int_{\{\mathbf{x}: x_1 > s_2 + m\}} \nu(d\mathbf{x})$  a.e.  $m$ . Applying this to  $\nu = Q \otimes H$  and  $\nu = Q' \otimes H$ , it follows that if  $u_{Q, \kappa_m} = u_{Q', \kappa_m}$  for all  $m \in \mathbb{R}$ , then  $\int_{\{\mathbf{x}: x_1 > s_2 + m\}} (Q \otimes H)(d(x_1, s_2)) = \int_{\{\mathbf{x}: x_1 > s_2 + m\}} (Q' \otimes H)(d(x_1, s_2))$  a.e.  $m$ . Rearranging, we have

$$\begin{aligned} \int_{S_2} (1 - Q(s_2 + m)) a(s_2) ds_2 &= \int_{S_2} (1 - Q'(s_2 + m)) a(s_2) ds_2 \\ \implies 0 &= \int_{S_2} (Q(m - s_2) - Q'(m - s_2)) a(s_2) ds_2, \end{aligned} \quad (9)$$

by the evenness of the normal distribution. Letting  $K := Q - Q'$ , we have that the convolution  $[K * f](m) = 0$  a.e.  $m$ . Let  $\mathcal{F}$  denote the normalized Fourier transform,  $\mathcal{F}(g)(z) := \int_{-\infty}^{\infty} e^{-2\pi izx} g(x) dx$ . Both  $K$  and  $h$  are Lebesgue integrable functions,<sup>31</sup> and thus by the convolution theorem,<sup>32</sup>  $\mathcal{F}(K * h) = \mathcal{F}(K) \cdot \mathcal{F}(h)$ . But since  $K * h \equiv 0$ ,  $\mathcal{F}(K * h) \equiv 0$ . Since  $\mathcal{F}(h)(z) = e^{-2(\pi\sigma z)^2} > 0$  for all  $z$ , we

<sup>27</sup>Here we use the fact that  $G_{\kappa_m}(\cdot; \theta)$  is constant in  $\theta$  for any constant markup strategy  $\kappa_m$ .

<sup>28</sup>This follows from the fact that  $(\min\{x_1, x_2\})^2 \leq x_1^2 + x_2^2$ , the bounds (2), and the finiteness of second moments for normal distributions.

<sup>29</sup>See Royden and Fitzpatrick (1988, Corollary 5.15).

<sup>30</sup>See Royden and Fitzpatrick (1988, Theorem 5.10).

<sup>31</sup>The integrability of  $H$  follows from the bounds (2) applied to  $Q$  and  $Q'$ , and the fact that for all  $M > 0$ ,  $\int_{-\infty}^{\infty} (H(x + M) - H(x - M)) dx = 2M < \infty$ .

<sup>32</sup>See Reed and Simon (1980, Theorem IX.3b).



must have  $\mathcal{F}(K) \equiv 0$ . Applying the inverse Fourier transform, we have  $K = 0$  almost everywhere. This contradicts the assumption that  $Q \neq Q'$ , so it must be that  $u_{Q, \kappa_m} \neq u_{Q', \kappa_m}$  for some  $m$ .

Next, we show that  $u_{Q, \kappa_m} \neq u_{Q', \kappa_m}$  contradicts weak admissibility. By continuity of  $G_\mu(\cdot, \theta)$  in  $\theta$ , for each  $\eta > 0$ , we there exists  $\delta > 0$  such that  $d_P(G_\mu(\cdot; \theta), Q) < \delta \implies |u(\theta) - u_{Q, \kappa_m}| < \eta$  and  $d_P(G_\mu(\cdot; \theta), Q') < \delta \implies |u(\theta) - u_{Q', \kappa_m}| < \eta$ .<sup>33</sup> Recall that by weak admissibility, there is some  $u^*$  such that  $\lim_{n \rightarrow \infty} \int u(\theta) d\mathbb{P}_n = u^*$  for all diffusing sequences  $(\mathbb{P}_n)$ . For the contradiction, we construct two sequence of measures  $(\mathbb{P}_n^1)$  and  $(\mathbb{P}_n^2)$  along which the limits differ. Since  $Q$  is an attraction and  $\lambda\{\theta : d_P(G_\mu(\cdot, \theta), Q) < \delta\} = \infty$ , for each  $n \in \mathbb{N}$ , there exists  $C_n^1 \subset \{\theta : d_P(G_\mu(\cdot, \theta), Q) < \delta\} \setminus [-n, n]$  with  $\lambda C_n^1 = 2n^2$ . Define  $B_n^1 := [-n, n] \cup C_n^1$ , and define  $\mathbb{P}_n^1(\theta) := \mathbf{1}_{B_n^1}(\theta) / \lambda(B_n^1)$ . By construction,  $(\mathbb{P}_n^1)_{n \in \mathbb{N}}$  is a diffusing sequence of measures, and by the assumption that  $\mu$  is weakly admissible, we must have  $\lim_{n \rightarrow \infty} \int u(\theta) d\mathbb{P}_n^1 = u^*$ . Now  $u(\theta)$  as a function of  $\theta$  is bounded in magnitude by some  $C > 0$ . We have  $|\int u(\theta) d\mathbb{P}_n^1 - u_{Q, \kappa_m}| \leq \int |u(\theta) - u_{Q, \kappa_m}| d\mathbb{P}_n^1 \leq \frac{\eta \cdot 2n^2 + 2C \cdot 2n}{2n^2 + 2n} \rightarrow \eta$  and thus  $u^* \in [u_{Q, \kappa_m} - \eta, u_{Q, \kappa_m} + \eta]$ . Likewise,  $Q'$  is an attraction, so  $\lambda\{\theta : d_P(G_\mu(\cdot, \theta), Q') < \delta\} = \infty$ , and we define  $C_n^2 \subset \{\theta : d_P(G_\mu(\cdot, \theta), Q') < \delta\} \setminus [-n, n]$  with  $\lambda C_n^2 = 2n^2$ ,  $B_n^2 = [-n, n] \cup C_n^2$ , and  $\mathbb{P}_n^2(\theta) := \mathbf{1}_{B_n^2}(\theta) / \lambda(B_n^2)$ , such that  $|\int u(\theta) d\mathbb{P}_n^2 - u_{Q', \kappa_m}| \rightarrow \eta$ , and thus  $u^* \in [u_{Q', \kappa_m} - \eta, u_{Q', \kappa_m} + \eta]$ . Since  $\eta$  is arbitrary, we choose  $\eta < \frac{|u_{Q, \kappa_m} - u_{Q', \kappa_m}|}{2}$  and obtain a contradiction. Hence we conclude that  $Q$  is unique.

In the final step, we prove that  $\mu$  is nearly stationary, and in particular  $\mu$  is nearly  $Q$ , by showing that for all  $\eta > 0$ ,  $\lambda\{\theta : d_P(G_\mu(\cdot, \theta), Q) \geq \eta\} < \infty$ . We derive a contradiction by showing that otherwise, the uniqueness result above would be violated. Let  $\mathcal{E}$  be the set of all such  $\eta$ ; by way of contradiction, we suppose  $\bar{\eta} = \inf\{\eta \in \mathcal{E}\} > 0$ .<sup>34</sup> Pick  $\eta^* \in (0, \bar{\eta})$  and let  $Y_{\eta^*} := \{Q \in \mathcal{G}^M : d_P(Q, Q) \geq \eta^*\}$ ; clearly  $Q_0 \notin Y_{\eta^*}$ . We have  $\lambda\{\theta : G_\mu(\cdot, \theta) \in Y_{\eta^*}\} = \infty$ , and  $Y_{\eta^*}$  is a compact subspace of  $\mathcal{G}^M$ . By Lemma 3, there exists an attraction  $Q' \in Y_{\eta^*}$ . This contradicts the uniqueness of the attraction  $Q$ , so  $\mu$  is near  $Q$ , completing the proof.  $\square$

## A.2 Proofs for Section 3.3

*Proof of Lemma 4.* To show that the mapping  $H_\mu$  is continuous in  $\theta$ , for any  $\theta$ , choose any  $x$  at which  $H_\mu(\epsilon, x; \theta)$  is continuous in  $(\epsilon, x)$ , and consider any sequence  $\theta_n \rightarrow \theta$ , which implies that  $\mu((-\infty, x]|s)h(s|\theta_n) \rightarrow \mu((-\infty, x]|s)h(s|\theta)$  pointwise in  $s$ . Since the  $\mu(\cdot|s)$  are probability measures, we have

$$\mu((-\infty, x]|s)h(s|\theta_n) \leq h(s|\theta_n).$$

Note that  $1 = \int_{-\infty}^{\infty} h(s_i|\theta_n)\lambda(ds_i) \rightarrow \int_{-\infty}^{\infty} h(s_i|\theta)\lambda(ds_i) = 1$ . By application of the generalized Lebesgue Convergence Theorem<sup>35</sup> and continuity of the integral,

$$\begin{aligned} H_\mu(\epsilon, x; \theta_n) &= \int_{-\infty}^{\epsilon + \theta_n} \mu((-\infty, x]|s)h(s|\theta_n)\lambda(ds) \\ &\rightarrow \int_{-\infty}^{\epsilon + \theta} \mu((-\infty, x]|s)h(s|\theta)\lambda(ds) = H_\mu(\epsilon, x; \theta). \end{aligned}$$

<sup>33</sup>Here we remind the reader that  $Q, Q', m$ , and  $\mu$  are fixed;  $\theta$  is arbitrary.

<sup>34</sup>The metric definition ensures that  $\eta^* \leq 1$ .

<sup>35</sup>See Royden and Fitzpatrick (1988, Theorem 4.17).

Hence  $H_\mu(\cdot, \cdot; \theta_n)$  converges weakly to  $H_\mu(\cdot, \cdot; \theta)$ , as desired.  $\square$

*Proof of Theorem 2.* Given an expert strategy  $\mu$ , we have  $\mu \in \mathcal{M}^M$  for some  $M \in \mathbb{N}$ . Suppose that  $\mu$  is not nearly stationary. By a repeat of arguments in the proof of Theorem 1, there exists an attraction  $Q \in \mathbb{H}^M$  for  $H_\mu$ .

To prove uniqueness, we suppose both  $Q$  and  $Q' \neq Q$  are attractions for  $\mu$ . Then there exists a pair  $(\epsilon, x)$  at which both  $Q'$  and  $Q$  are continuous but  $Q'(\epsilon, x) \neq Q(\epsilon, x)$ . As in the proof of Theorem 1, we can construct two sequence of measures  $(\mathbb{P}_n^1)$  and  $(\mathbb{P}_n^2)$  along which the limits converge to neighborhoods of  $Q(\epsilon, x)$  and  $Q'(\epsilon, x)$ , respectively. This contradicts strong admissibility. Since the notational details are almost identical to those in the proof of Theorem 1, we do not repeat them here. The proof that  $\mu$  is near  $Q$  is also a repeat of previous arguments and is therefore omitted.

Next, we show that  $Q$  is induced by some mixture of constant markup strategies. To do this, we show that there exists a regular conditional probability for  $Q$ , interpreted as a CDF  $F_{Q,X}(\cdot|\epsilon)$  over markups conditional on  $\epsilon$ , which is constant in  $\epsilon$ . Suppose otherwise that for any such  $F_{Q,X}(\cdot|\cdot)$ , there exists an  $x \in X_i^M$  and  $\epsilon, \epsilon' \in S_i$  such that  $F_{Q,X}(x|t)$  is continuous in  $x$  at  $x$  for  $t = \epsilon, \epsilon'$  but  $F_{Q,X}(x|\epsilon) \neq F_{Q,X}(x|\epsilon')$ . Since  $\mu$  is near  $Q$ , for any diffusing sequence  $(\mathbb{P}_n)$ , we have

$$\int_{-\infty}^{\infty} \mu(x|\epsilon + \theta) d\mathbb{P}_n(\theta) \rightarrow F_{Q,X}(x|\epsilon).$$

Now consider a change of variables  $\theta' = \theta + \epsilon - \epsilon'$ :

$$\int_{-\infty}^{\infty} \mu(x|\theta' + \epsilon') d\mathbb{P}_n(\theta') \rightarrow F_{Q,X}(x|\epsilon).$$

But  $\int_{-\infty}^{\infty} \mu(x|\theta' + \epsilon') d\mathbb{P}_n(\theta') \rightarrow F_{Q,X}(x|\epsilon')$ , so  $F_{Q,X}(x|\epsilon) = F_{Q,X}(x|\epsilon')$ , a contradiction. Given the existence of this  $F_{Q,X}(\cdot|\cdot)$  constant in  $\epsilon$ , we define a CDF over markups  $K_Q$  by  $K_Q(x) := F_{Q,X}(x|\epsilon)$  for all  $x \in X_i^M, \epsilon \in S_i$ .

The last step of the proof is to show payoff equivalence between  $\mu$  and  $Q$  (and hence  $K_Q$ ). Take any bias  $b$ , strongly admissible rival expert strategy  $\mu_2$ , and principal strategy  $C$ , and fixed  $\theta$ . Let the variables  $Q', Q'_2 \in \mathcal{H}^M$  denote an arbitrary joint distribution over noise and markups. Define the expression

$$u(Q', Q'_2) := \int \int [-\mathbb{1}\{C(a_1, a_2) = 1\}(a_1 - b)^2 - \mathbb{1}\{C(a_1, a_2) = 2\}(a_2 - b)^2] dQ'(\epsilon_2, x_2) dQ'(\epsilon_1, x_1), \quad (10)$$

where  $a_i = \epsilon_i + x_i$ .<sup>36</sup> Note that by familiar convergence arguments, the function  $u$  in (10) is continuous in its arguments and bounded. Evaluating  $u$  at  $Q' = H_\mu(\cdot, \cdot; \theta)$ ,  $Q'_2 = H_{\mu_2}(\cdot, \cdot; \theta)$  gives the expected payoff conditional on  $\theta$  for expert strategy profile  $(\mu, \mu_2)$ ; we denote this as  $u(\theta) := u(H_\mu(\cdot, \cdot; \theta), H_{\mu_2}(\cdot, \cdot; \theta))$  with some abuse of notation. Now  $\mu$  is near  $Q$  and if  $\mu_2$  is strongly admissible, it is near some  $Q_2$ . Let  $u^* := u(Q, Q_2)$ . Now for any  $\eta > 0$ , by continuity there exists  $\delta > 0$  such that if  $\max\{d_P(H_\mu(\cdot, \cdot; \theta), Q), d_P(H_{\mu_2}(\cdot, \cdot; \theta), Q_2)\} < \delta$ , then  $|u(\theta) - u^*| < \eta$ . By nearness

<sup>36</sup>Since the principal is restricted to simple strategies, the choice is translation invariant, and we thus drop  $\theta$ .

to  $Q$  and  $Q_2$ ,  $\max\{d_P(H_\mu(\cdot, \cdot; \theta), Q), d_P(H_{\mu_2}(\cdot, \cdot; \theta), Q_2)\} < \delta$  holds for all but a finite measure of  $\theta$ . Thus for any  $\eta > 0$ ,  $|u(\theta) - u^*| < \eta$  for all but a finite measure of  $\theta$ . Integrating over any diffusing sequence  $(\mathbb{P}_n)$ , we have  $\int u(\theta) d\mathbb{P}_n(\theta) \rightarrow u^*$ . The same logic shows that  $\int u(Q, H_{\mu_2}(\cdot, \cdot; \theta)) d\mathbb{P}_n(\theta) \rightarrow u^*$ , establishing payoff equivalence between  $\mu$  and  $Q$ .  $\square$

### A.3 Proofs for Section 4

We first provide an auxiliary lemma, which is proved in Supplementary Appendix.

**Lemma A.1.** *Let  $\epsilon_1, \epsilon_2 \sim N(0, \sigma^2)$ , where  $\epsilon_1$  and  $\epsilon_2$  are independent, and define*

$$\xi(k_1, k_2) := \min(\epsilon_1 + k_1, \epsilon_2 + k_2), \quad \eta(k_1, k_2) := \max(\epsilon_1 + k_1, \epsilon_2 + k_2). \quad \text{Then}$$

$$\begin{aligned} \mathbb{E}\xi(k_1, k_2) &= -2\sigma^2 f(k_1 - k_2) + k_1(1 - F(k_1 - k_2)) + k_2 F(k_1 - k_2); \\ \mathbb{E}\eta(k_1, k_2) &= 2\sigma^2 f(k_1 - k_2) + k_1 F(k_1 - k_2) + k_2(1 - F(k_1 - k_2)); \\ \mathbb{E}\xi^2(k_1, k_2) &= \sigma^2 - 2(k_1 + k_2)\sigma^2 f(k_1 - k_2) + k_1^2(1 - F(k_1 - k_2)) + k_2^2 F(k_1 - k_2); \\ \mathbb{E}\eta^2(k_1, k_2) &= \sigma^2 + 2(k_1 + k_2)\sigma^2 f(k_1 - k_2) + k_1^2 F(k_1 - k_2) + k_2^2(1 - F(k_1 - k_2)). \end{aligned}$$

*Proof of Proposition 2.* After observing the signal  $s_i$ , expert  $i$  does a Bayesian update of his beliefs:

$$\theta|s_i \sim N(s_i, \sigma^2) \quad \text{and} \quad s_j|s_i \sim N(s_i, 2\sigma^2).$$

Since the principal chooses the lower offer, she accepts  $a_i$  iff  $s_j > s_i + k_i - k_j$ . Denote by  $g$  the PDF of  $N(s_i, 2\sigma^2)$ . Hence, the expected utility of expert  $i$

$$\begin{aligned} U_i(k_1, k_2, L) &= \int_{s_i + k_i - k_j}^{\infty} \mathbb{E}[B - (a_i - \theta - b_i)^2 | s_i, s_j] g(s_j) ds_j \\ &\quad + \int_{-\infty}^{s_i + k_i - k_j} \mathbb{E}[-(a_j - \theta - b_i)^2 | s_i, s_j] g(s_j) ds_j \end{aligned}$$

As  $(\theta|s_i, s_j) \sim N(\frac{s_j + s_i}{2}, \frac{\sigma^2}{2})$ ,  $a_i = s_i + k_i$ ,  $a_j = s_j + k_j$ , we obtain

$$\begin{aligned} U_i(k_1, k_2, L) &= \int_{s_i + k_i - k_j}^{\infty} \left[ B - \left( k_i - b_i - \frac{s_j - s_i}{2} \right)^2 - \frac{\sigma^2}{2} \right] g(s_j) ds_j \\ &\quad + \int_{-\infty}^{s_i + k_i - k_j} \left[ - \left( k_j - b_i + \frac{s_j - s_i}{2} \right)^2 - \frac{\sigma^2}{2} \right] g(s_j) ds_j. \end{aligned}$$

Now make a substitution  $t = s_j - s_i$  and denote by  $f$  and  $F$  the PDF and CDF of  $N(0, 2\sigma^2)$ .

$$U_i(k_1, k_2, L) = \int_{k_i - k_j}^{\infty} \left[ B - \left( k_i - b_i - \frac{t}{2} \right)^2 - \frac{\sigma^2}{2} \right] f(t) dt + \int_{-\infty}^{k_i - k_j} \left[ - \left( k_j - b_i + \frac{t}{2} \right)^2 - \frac{\sigma^2}{2} \right] f(t) dt.$$

Note that  $U_i(k_1, k_2, L)$  does not depend on signal  $s_i$ , which is intuitive for the improper prior. As  $\int_a^{\infty} t f(t) dt = 2\sigma^2 f(a)$  and  $\int_{-\infty}^{\infty} t^2 f(t) dt = 2\sigma^2$ , we get the expression for  $U_i(k_i, k_j, L)$ .

Now in state  $\theta$ , the principal's action  $a$  is distributed as  $\theta + \xi$ , where  $\xi = \min(\epsilon_1 + k_1, \epsilon_2 + k_2)$ ;  $\epsilon_1, \epsilon_2 \sim N(0, \sigma^2)$ ,  $\epsilon_1$  and  $\epsilon_2$  are independent.

Therefore, from Lemma A.1 the expected bias of the accepted offer is

$$b(k_1, k_2, L) = \mathbb{E}\xi(k_1, k_2) = -2\sigma^2 f(k_1 - k_2) + k_2 F(k_1 - k_2) + k_1(1 - F(k_1 - k_2))$$

and the expected utility of the principal is

$$\begin{aligned} V(k_1, k_2, L) &= -\mathbb{E}(a - \theta)^2 = -\mathbb{E}(\theta + \xi - \theta)^2 = -\mathbb{E}\xi^2(k_1, k_2) \\ &= -\sigma^2 + 2(k_1 + k_2)\sigma^2 f(k_1 - k_2) - k_2^2 - (k_1^2 - k_2^2)(1 - F(k_1 - k_2)). \end{aligned}$$

Finally, the variance of the chosen offer is

$$\begin{aligned} \text{Var}(k_1, k_2, L) &= -V(k_1, k_2, L) - b^2(k_1, k_2, L) \\ &= \sigma^2 - 4\sigma^4 f^2(z) - 2\sigma^2 z f(z)(2F(z) - 1) + z^2 F(z)(1 - F(z)). \end{aligned}$$

□

*Proof of Theorem 4.* We start by showing that  $U_i(k_1, k_2, L)$  is a single-peaked function of  $k_i$ . Taking a derivative w.r.t.  $k_i$  yields

$$\begin{aligned} U'_i(k_i) &= -2[(k_i - b_i)(1 - F(k_i - k_j)) - \rho(k_i - k_j)] \\ &= -2(1 - F(k_i - k_j)) \left[ k_i - b_i - \rho \frac{f(k_i - k_j)}{1 - F(k_i - k_j)} \right]. \end{aligned}$$

Let  $g(k_i)$  denote the term in square brackets above. Lemma 6 implies that  $g'(k_i) = 1 - \rho\lambda(k_i - k_j) \geq 1 - \sigma^2\lambda(k_i - k_j) > 0$ . Additionally, we have  $\lim_{x \rightarrow \pm\infty} g(x) = \pm\infty$ . Combining these facts,  $U_i$  has a unique critical point, which is a global maximum.

We now look for upward equilibria. The FOCs for the experts are equivalent to:

$$k_1 - b_1 - \rho \frac{f(k_1 - k_2)}{1 - F(k_1 - k_2)} = 0 \tag{11}$$

$$k_2 - b_2 - \rho \frac{f(k_1 - k_2)}{F(k_1 - k_2)} = 0. \tag{12}$$

Subtracting (12) from (11) substituting  $z = k_1 - k_2$ , we get

$$z - \rho \left[ \frac{f(z)}{1 - F(z)} - \frac{f(z)}{F(z)} \right] = b_1 - b_2. \tag{13}$$

Denote  $l(z) = -\rho \left[ \frac{f(z)}{1 - F(z)} - \frac{f(z)}{F(z)} \right] + z$ . Using a) from Lemma 6, we obtain

$$l'(z) = 1 - \rho[v'(z) - w'(z)] + 1 \geq 1 - \sigma^2[v'(z) - w'(z)] > 0$$

Now  $l(z)$  is continuous, strictly increasing on  $\mathbb{R}$ , and ranges from  $-\infty$  to  $+\infty$ . Therefore (4) has a unique solution,  $z^*$ ; we use  $z(B)$  to denote explicitly the dependence on  $B$ .

Using this solution, we get  $(k_1^U, k_2^U)$  as the only critical point and check that this point satisfies both initial FOCs. As  $U_i(k_i, k_j, L)$  is a single-peaked function of  $k_i$ ,  $(k_1^U, k_2^U)$  is a pair of best responses.

As it was shown in Theorem 3, choosing the lower offer is the BR strategy for the principal iff  $k_1 + k_2 \geq 0$ , or equivalently

$$b_1 + b_2 - \left[ \frac{f(z^*)}{1 - F(z^*)} + \frac{f(z^*)}{F(z^*)} \right] \geq 0.$$

Also the LHS of (4) is equal to 0 at  $z = 0$ , and therefore  $z^* \geq 0$  and  $k_1^U - k_2^U \geq 0$ .

Define a function  $m(B) = b_1 + b_2 + \rho[v(z(B)) + w(z(B))]$ ; the upward equilibrium exists if and only if  $m(B) \geq 0$ .

1) For  $B \leq 2\sigma^2$ :  $m(B) \geq 0$ , therefore the upward equilibrium exists.

2) Next, we show that  $m(B)$  is decreasing in  $B$  in the region  $B \geq 2\sigma^2$ .

$$m'(B) = -\frac{1}{2}[v(z(B)) + w(z(B))] + \rho[\lambda(z(B)) - \lambda(-z(B))]z'(B). \quad (14)$$

Differentiating equation (4) at point  $B$ , we get:

$$z'(B) - \rho[\lambda(z(B)) + \lambda(-z(B))]z'(B) + \frac{1}{2}[v(z(B)) - w(z(B))] = 0. \quad (15)$$

By substituting (15), the second term of (14) becomes

$$\begin{aligned} & + \rho[\lambda(z(B)) - \lambda(-z(B))] \frac{-\frac{1}{2}[v(z(B)) - w(z(B))]}{1 - \rho[\lambda(z(B)) + \lambda(-z(B))]} \\ & \leq -\rho[\lambda(z(B)) - \lambda(-z(B))] \frac{\frac{1}{2}[v(z(B)) - w(z(B))]}{-\rho[\lambda(z(B)) + \lambda(-z(B))]} \\ & = \frac{1}{2} \frac{\lambda(z(B)) - \lambda(-z(B))}{\lambda(z(B)) + \lambda(-z(B))} [v(z(B)) - w(z(B))] \\ \implies m'(B) & \leq -\frac{1}{2}[v(z(B)) + w(z(B))] + \frac{1}{2}[v(z(B)) - w(z(B))] = -w(z(B)) < 0. \end{aligned}$$

3) From Lemma 6 the hazard rate  $v$  is convex, so for any real  $x$ ,  $v(x) + w(x) = v(x) + v(-x) \geq 2v(0) > 0$ , and  $m(B)$  tends to  $-\infty$  as  $B$  tends to  $\infty$ .

From 1)-3) follows that there exists  $B_U : m(B) \geq 0$  iff  $B \leq B_U$ . Also

$$\left( \frac{B_U}{2} - \sigma^2 \right) [v(z(B_U)) + w(z(B_U))] = b_1 + b_2. \quad (16)$$

As  $z(B_U)$  satisfies equation (4), we have:

$$\left( \frac{B_U}{2} - \sigma^2 \right) [v(z(B_U)) - w(z(B_U))] + z(B_U) = b_1 - b_2. \quad (17)$$

rom the previous discussion and (16) we have  $B_U \geq 2\sigma^2$ . Also, (16) and the inequality  $v(x) + w(x) = v(x) + v(-x) \geq 2v(0) = \frac{2}{\sqrt{\pi}\sigma}$  give an upper bound on  $B_U$ :

$$\left(\frac{B_U}{2} - \sigma^2\right) \frac{2}{\sqrt{\pi}\sigma} \leq b_1 + b_2.$$

Subtracting (17) from (16), we get a lower bound on  $B_U$ :

$$2b_2 = (B_U - 2\sigma^2)w(z(B_U)) - z(B_U) \leq (B_U - 2\sigma^2)w(0) = (B_U - 2\sigma^2) \frac{1}{\sqrt{\pi}\sigma}.$$

Finally, we calculate the expected bias of the chosen offer, its variance, and players' utilities:

$$\begin{aligned} b(k_1^U, k_2^U, L) &= -2\sigma^2 f(z^*) + k_2^U F(z^*) + k_1^U (1 - F(z^*)) \\ &= -2\sigma^2 f(z^*) + b_2 F(z^*) + \rho f(z^*) + b_1 (1 - F(z^*)) + \rho f(z^*) \\ &= b_1 (1 - F(z^*)) + b_2 F(z^*) - B f(z^*); \\ \text{Var}(k_1^U, k_2^U, L) &= \sigma^2 - 4\sigma^4 f^2(z^*) - 2\sigma^2 z^* f(z^*) (2F(z^*) - 1) + (z^*)^2 F(z^*) (1 - F(z^*)). \end{aligned}$$

□

**Corollary A.1.** *In upward equilibrium,*

- $V(k_1^U, k_2^U, L) = -\sigma^2 - b_1^2(1 - F(z^*)) - b_2^2 F(z^*) + B(b_1 + b_2)f(z^*) + \left(\sigma^4 - \frac{B^2}{4}\right) \frac{f^2(z^*)}{F(z^*)(1 - F(z^*))}$ ;
- $U_1(k_1^U, k_2^U, L) = -\sigma^2 - (b_1 - b_2)^2 F(z^*) + B(1 - F(z^*)) - B(b_1 - b_2)f(z^*) + \left(\sigma^4 - \frac{B^2}{4}\right) \frac{f^2(z^*)}{F(z^*)(1 - F(z^*))}$ ;
- $U_2(k_1^U, k_2^U, L) = -\sigma^2 - (b_1 - b_2)^2 (1 - F(z^*)) + B F(z^*) + B(b_1 - b_2)f(z^*) + \left(\sigma^4 - \frac{B^2}{4}\right) \frac{f^2(z^*)}{F(z^*)(1 - F(z^*))}$ .

*Proof.* Immediate from applying Proposition 2 to the markups given by Theorem 4. □

*Proof of Proposition 3.* If  $b_1 = b_2 = b > 0$ , then upper and lower bounds on  $B_U$  coincide, and thus  $B_U = 2\sigma^2 + 2\sqrt{\pi}\sigma b$ . From Theorem 4, the experts' markups are  $k_1^U = k_2^U = k_U = b + \frac{\rho\sigma}{\sigma\sqrt{\pi}}$  and  $z^* = 0$ . The other results follow immediately. □

*Proof of Proposition 4.* Since the principal chooses the highest offer, she chooses  $a_i$  iff  $s_j < s_i + k_i - k_j$ . Using arguments similar to used in Proposition 2, we find the expected utility of expert  $i$ :

$$\begin{aligned} U_i(k_1, k_2, H) &= \int_{k_i - k_j}^{\infty} \mathbb{E}[-(a_j - \theta - b_i)^2 | s_j] f(s_j) ds_j + \int_{-\infty}^{k_i - k_j} \mathbb{E}[B - (a_i - \theta - b_i)^2 | s_j] f(s_j) ds_j \\ &= \int_{k_i - k_j}^{\infty} \left[ -\left(k_j - b_i + \frac{s_j}{2}\right)^2 - \frac{\sigma^2}{2} \right] f(s_j) ds_j + \int_{-\infty}^{k_i - k_j} \left[ B - \left(k_i - b_i - \frac{s_j}{2}\right)^2 - \frac{\sigma^2}{2} \right] f(s_j) ds_j \\ &= (B - (k_i - b_i)^2) F(k_i - k_j) - \sigma^2 - (k_j - b_i)^2 [1 - F(k_i - k_j)] - 2\sigma^2 (k_i + k_j - 2b_i) f(k_i - k_j). \end{aligned}$$

In state  $\theta$  the principal's action  $a$  is distributed as  $\theta + \eta$ , where  $\eta \sim \max(\epsilon_1 + k_1, \epsilon_2 + k_2)$ ;  $\epsilon_1, \epsilon_2 \sim N(0, \sigma^2)$ ,  $\epsilon_1$  and  $\epsilon_2$  are independent.

From Lemma A.1 the expected bias of the accepted offer is

$$b(k_1, k_2, H) = \mathbb{E}\eta(k_1, k_2) = 2\sigma^2 f(k_1 - k_2) + k_2(1 - F(k_1 - k_2)) + k_1 F(k_1 - k_2).$$

The expected utility of the principal is

$$\begin{aligned} V(k_1, k_2, H) &= -\mathbb{E}(a - \theta)^2 = -\mathbb{E}\eta^2(k_1, k_2) \\ &= -\sigma^2 - 2(k_1 + k_2)\sigma^2 f(k_1 - k_2) - k_2^2 - (k_1^2 - k_2^2)F(k_1 - k_2). \end{aligned}$$

The variance of the chosen offer is

$$\begin{aligned} Var(k_1, k_2, H) &= -V(k_1, k_2, H) - b^2(k_1, k_2, H) \\ &= \sigma^2 - 4\sigma^4 f^2(z) - 2\sigma^2 z f(z)(2F(z) - 1) + z^2 F(z)(1 - F(z)). \end{aligned}$$

□

*Proof of Theorem 5.* The proof is analogous to that of Theorem 4. The FOCs for experts are now:

$$k_1 - b_1 + \rho \frac{f(k_1 - k_2)}{F(k_1 - k_2)} = 0 \quad (18)$$

$$k_2 - b_2 + \rho \frac{f(k_1 - k_2)}{1 - F(k_1 - k_2)} = 0. \quad (19)$$

Subtracting equation (18) from equation (19) yields (13). Principal optimality holds if and only if  $k_1 + k_2 \leq 0$ , or equivalently

$$n(B) := b_1 + b_2 + \left[ \frac{f(z^*)}{1 - F(z^*)} + \frac{f(z^*)}{F(z^*)} \right] \leq 0, \quad (20)$$

where  $z(B)$  is given by equation (4). For  $B > 2\sigma^2$ , we have  $n(B) > 0$ , and thus a downward equilibrium does not exist. Observe further that  $n(2\sigma^2) = b_1 + b_2 \geq 0$ . Since  $m(B) + n(B) = 2(b_1 + b_2)$  and  $m'(B) < 0$ , we have  $n'(B) > 0$ . It follows that if  $n(0) \leq 0$ , then there exists  $B_D \in [0, 2\sigma^2]$  such that  $n(B) \leq 0$  iff  $B \leq B_D$ . Therefore

$$\left( \frac{B_U}{2} - \sigma^2 \right) [v(z(B_D)) + w(z(B_D))] = -(b_1 + b_2). \quad (21)$$

Also  $z(B_D)$  satisfies equation (4), and therefore

$$\left( \frac{B_U}{2} - \sigma^2 \right) [v(z(B_D)) - w(z(B_D))] + z(B_D) = b_1 - b_2. \quad (22)$$

From the previous discussion and (21) we have  $B_D \leq 2\sigma^2$ . Also (21) and the inequality  $v(x) + w(x) \geq$



$2v(0) = \frac{2}{\sqrt{\pi}\sigma}$  give the lower bound

$$\left(\frac{B_D}{2} - \sigma^2\right) \frac{2}{\sqrt{\pi}\sigma} \geq -(b_1 + b_2).$$

Summing (22) and (21), we get the upper bound

$$-2b_2 = \left(\frac{B_U}{2} - \sigma^2\right) v(z(B_D)) + z(B_D) \geq (B_D - 2\sigma^2)2v(0).$$

Finally, we compute the following:

$$\begin{aligned} b(k_1^D, k_2^D, H) &= 2\sigma^2 f(z^*) + k_1^D F(z^*) + k_2^D (1 - F(z^*)) \\ &= 2\sigma^2 f(z^*) + b_1 F(z^*) - \rho f(z^*) + b_2 (1 - F(z^*)) - \rho f(z^*) \\ &= b_1 F(z^*) + b_2 (1 - F(z^*)) + B f(z^*); \\ \text{Var}(k_1^D, k_2^D, H) &= \sigma^2 - 4\sigma^4 f^2(z^*) - 2\sigma^2 z^* f(z^*) (2F(z^*) - 1) + (z^*)^2 F(z^*) (1 - F(z^*)). \end{aligned}$$

□

**Corollary A.2.** *In downward equilibrium,*

$$\begin{aligned} V(k_1^D, k_2^D, H) &= -\sigma^2 - b_1^2 F(z^*) - b_2^2 (1 - F(z^*)) - B(b_1 + b_2) f(z^*) + \left(\sigma^4 - \frac{B^2}{4}\right) \frac{f^2(z^*)}{F(z^*)(1 - F(z^*))}; \\ U_1(k_1^D, k_2^D, H) &= -\sigma^2 - (b_1 - b_2)^2 (1 - F(z^*)) + B F(z^*) + B(b_1 - b_2) f(z^*) + \left(\sigma^4 - \frac{B^2}{4}\right) \frac{f^2(z^*)}{F(z^*)(1 - F(z^*))}; \\ U_2(k_1^D, k_2^D, H) &= -\sigma^2 - (b_1 - b_2)^2 F(z^*) + B(1 - F(z^*)) - B(b_1 - b_2) f(z^*) + \left(\sigma^4 - \frac{B^2}{4}\right) \frac{f^2(z^*)}{F(z^*)(1 - F(z^*))}. \end{aligned}$$

*Proof.* Proposition 4 applied to Theorem 5. □

*Proof of Proposition 5.* If  $b_1 = b_2 = b > 0$ , then upper and lower bounds on  $B_D$  from Theorem 5 coincide, so  $B_D = 2\sigma^2 + 2\sqrt{\pi}\sigma b$ . The experts' markups are  $k_1^D = k_2^D = k_D = b - \frac{\rho}{\sigma\sqrt{\pi}}$  and  $z^* = 0$ , which implies all other results. □

## A.4 Proofs for Section 5

*Proof of Proposition 6.* From Theorems 4 and 5

$$V_{upw.} - V_{downw.} = (2F(z^*) - 1)(b_1^2 - b_2^2) + 2B(b_1 + b_2)f(z^*) \geq 0,$$

with equality if and only if either  $b_1 + b_2 = 0$  or both  $B = 0$  and  $b_1 = b_2$ . □

*Proof of Proposition 9.* From Proposition 3, the symmetric upward equilibrium exists if  $B \leq 2\sigma^2 + 2\sqrt{\pi}b\sigma$  or, equivalently,  $\sigma \geq \frac{\sqrt{\pi b^2 + 2B} - \sqrt{\pi}b}{2}$ . The principal's expected payoff in upward equilibrium is equal to  $V = -(b - \frac{B}{2\sqrt{\pi}\sigma})^2 - (1 - \frac{1}{\pi})\sigma^2$ . Then  $V'(\sigma) = \frac{B}{\sqrt{\pi}\sigma^2}(\frac{B}{2\sqrt{\pi}\sigma} - b) - 2(1 - \frac{1}{\pi})\sigma = -\frac{1}{2\pi\sigma^3}[(4\pi - 4)\sigma^4 + 2\sqrt{\pi}bB\sigma - B^2] > 0$  if and only if  $\sigma < \sigma^*$ .

If  $B = 0$ ,  $V'(\sigma) < 0$ . Otherwise, denote  $\sigma_0 = \frac{\sqrt{\pi b^2 + 2B} - \sqrt{\pi}b}{2} > 0$ .

The interval  $[\sigma_0, \sigma^*)$  is non-empty if and only if  $0 > (4\pi - 4)\sigma_0^4 + 2\sqrt{\pi}bB\sigma_0 - B^2 = (4\pi - 4)\sigma_0^4 - 2B\sigma_0^2 + B[2\sigma_0^2 + 2\sqrt{\pi}b\sigma_0 - B] = (4\pi - 4)\sigma_0^4 - 2B\sigma_0^2$  or, equivalently,  $0 > (2\pi - 2)\sigma_0^2 - B = (2\pi - 2)\sigma_0^2 - (2\sigma_0^2 + 2\sqrt{\pi}b\sigma_0) = (2\pi - 4)\sigma_0^2 - 2\sqrt{\pi}b\sigma_0$ . The latter holds if and only if  $\sigma_0 < \frac{\sqrt{\pi}b}{\pi - 2}$  or, equivalently,  $B = 2\sigma_0^2 + 2\sqrt{\pi}b\sigma_0 < \frac{2(\pi - 1)\pi}{(\pi - 2)^2}b^2$ .  $\square$

*Proof of Proposition 10.* By Proposition 6, only upward equilibrium should be considered. Therefore, we seek to maximize the quadratic function  $V(k_U, k_U, L) - B = -\left(b - \frac{B}{2\sqrt{\pi}\sigma}\right)^2 - \sigma^2 + \frac{\sigma^2}{\pi} - B$  on the interval  $B \in [0, 2\sigma^2 + 2\sqrt{\pi}\sigma b]$ .

a) First consider  $\frac{b}{\sigma} \leq \sqrt{\pi}$ . In this case the maximum is achieved at  $B = 0$  and is equal to  $V(k_U, k_U, L|B = 0) = -b^2 - (1 - \frac{1}{\pi})\sigma^2$ .

Hence, if  $\frac{b}{\sigma} < \sqrt{\pi}$ , then the principal pays no bonus.

b) Now consider  $\frac{b}{\sigma} \geq \sqrt{\pi}$ . In this case the principal achieves maximum in upward equilibrium at  $B_R = 2\sqrt{\pi}\sigma(b - \sqrt{\pi}\sigma)$  (upward equilibrium exists for this point) and is equal to  $V(k_U, k_U, L|B = B_R) = (\pi - 1 + \frac{1}{\pi})\sigma^2 - 2\sqrt{\pi}\sigma b$ .

Her gains comparatively to  $B = 0$  (if she is legally restricted from paying bonuses) are equal to:

$$V(k_U, k_U, L|B = B_R) - V(k_U, k_U, L|B = 0) = (b - \sqrt{\pi}\sigma)^2.$$

$\square$

## A.5 Proofs for Section 6

*Proof of Proposition 12.* First, we calculate marginal utilities:

$$U'_i(k_i) = -2(k_i - b)[1 - F(k_i - k_j)] + \left[\frac{1}{2}\sigma^2 + 2\rho + \frac{1}{4}(k_i + k_j - 2b)^2\right] f(k_i - k_j).$$

Here, setting  $U'_i(k) = 0$  gives two critical points:

$$k = b + \sqrt{\pi}\sigma - \sqrt{\left(\pi - \frac{5}{2}\right)\sigma^2 + B} \text{ and } k = b + \sqrt{\pi}\sigma + \sqrt{\left(\pi - \frac{5}{2}\right)\sigma^2 + B}.$$

The second derivative is:

$$U''_i(k_i) = -2[1 - F(k_i - k_j)] + \left[2(k_i - b) + \frac{1}{2}(k_i + k_j - 2b) + \left(\frac{B}{2\sigma^2} - \frac{5}{4}\right)(k_i - k_j) - \frac{1}{8\sigma^2}(k_i - k_j)(k_i + k_j - 2b)^2\right] f(k_i - k_j).$$

We get that only  $k^* = b + \sqrt{\pi}\sigma - \sqrt{\left(\pi - \frac{5}{2}\right)\sigma^2 + B}$  is a local maximum of experts' utility functions.

Optimality for the principal holds if and only if  $k^* \geq 0$  or, equivalently,  $B \leq b^2 + 2\sqrt{\pi}b\sigma + \frac{5}{2}\sigma^2$ .

Calculating, we get that  $U_i(k^*, k^*, L) = \frac{(\pi - 1)\sigma}{\sqrt{\pi}}\sqrt{\left(\pi - \frac{5}{2}\right)\sigma^2 + B} - (\pi - \frac{7}{4})\sigma^2$ . As we noted earlier, a necessary condition for equilibrium is  $U_i(k^*, k^*, L) \geq 0$  or, equivalently,  $B \geq \left(\frac{5}{2} - \frac{3\pi(8\pi - 11)}{16(\pi - 1)^2}\right)\sigma^2$ .

Therefore, upward equilibrium may exist only if  $B \in \left[\left(\frac{5}{2} - \frac{3\pi(8\pi - 11)}{16(\pi - 1)^2}\right)\sigma^2, \frac{5}{2}\sigma^2 + 2\sqrt{\pi}b\sigma + b^2\right]$ .

To finish the proof, we show that if  $B$  lies on this interval, then  $k = k^*$  is a global maximum of  $U_1(k, k^*, L)$ .

Denote  $g(k) = -2(k-b) + \left[\frac{\sigma^2}{2} + 2\rho + \frac{1}{4}(k+k^* - 2b)^2\right] v(k-k^*)$ .<sup>37</sup>

Then  $U_1'(k) = -2(k-b)[1-F(k-k^*)] + \left[\frac{\sigma^2}{2} + 2\rho + \frac{1}{4}(k+k^* - 2b)^2\right] f(k-k^*) = [1-F(k-k^*)]g(k)$  and  $\text{sign}(U_1'(k)) = \text{sign}(g(k))$ .

The first and second derivatives of  $g$  are

$$g'(k) = -2 + \left[\frac{\sigma^2}{2} + 2\rho + \frac{1}{4}(k+k^* - 2b)^2\right] v'(k-k^*) + \frac{1}{2}(k+k^* - 2b)v(k-k^*)$$

$$g''(k) = \left[\frac{\sigma^2}{2} + 2\rho - B + \frac{1}{4}(k+k^* - 2b)^2\right] v''(k-k^*) + (k+k^* - 2b)v'(k-k^*) + \frac{1}{2}v(k-k^*).$$

Consider two cases.

1.  $B \in \left[\left(\frac{5}{2} - \frac{3\pi(8\pi-11)}{16(\pi-1)^2}\right)\sigma^2, \frac{5}{2}\sigma^2\right]$ . Here  $k^* \geq b$ .

a) On the interval  $k < b$ ,  $U_1'(k) > 0$  and hence there is no point of maximum there.

b) On the interval  $k \geq b$  we also have that  $k+k^* - 2b \geq 0$ . As all  $v$ ,  $v'$  and  $v''$  are strictly positive functions,  $g''(k) > 0$ . As  $g'(k^*) < 0$  and  $g'(+\infty) > 0$ , hence there exists  $k^{**} > k^*$ : for  $k < k^{**}$   $g(k)$  is decreasing; for  $k > k^{**}$ ,  $g(k)$  is increasing. As  $g(b) > 0$ ,  $g(k^*) = 0$ ,  $g(k^{**}) < 0$  and  $g(+\infty) > 0$ , then there exists  $k_0 > k^{**}$ :  $g(k_0) = 0$ . In summary,  $g(k)$  is negative only on  $(k^*, k_0)$ . Consequently,  $U_1(k)$  is increasing on  $[b, k^*)$ , decreasing on  $(k^*, k_0)$ , increasing for  $k > k_0$ . Hence, to show that  $k^*$  is a maximum on the interval  $k \geq b$  it is sufficient to verify that  $U_1(k^*) \geq U_1(+\infty) = 0$ , which was already done.

2.  $B \in \left[\frac{5}{2}\sigma^2, \frac{5}{2}\sigma^2 + 2\sqrt{\pi}b\sigma + b^2\right]$ . Here  $k^* \leq b$ .

a) On the interval  $k < k^*$ :  $U_1'(k) > 2(b-k^*)[1-F(k^*-k^*)] + \left[\frac{\sigma^2}{2} + 2\rho + \frac{1}{4}(k^*+k^* - 2b)^2\right] f(k-k^*) = b-k^* - \frac{b-k^*}{f(0)}f(k-k^*)$  (as  $k^*$  is a solution of  $\frac{5}{2}\sigma^2 - B + (k^* - b)^2 = \frac{k^*-b}{f(0)}$ ). Therefore  $U_1'(k) > b-k^* - \frac{b-k^*}{f(0)}f(k-k^*) \geq b-k^* - \frac{b-k^*}{f(0)}f(0) = 0$ .

b) On the interval  $k \in (k^*, b]$ :  $\frac{U_1'(k)}{f(k-k^*)} = -2(k-b)\frac{1-F(k-k^*)}{f(k-k^*)} + \left[\frac{\sigma^2}{2} + 2\rho + \frac{1}{4}(k+k^* - 2b)^2\right] < 2(b-k^*)\frac{1-F(k^*-k^*)}{f(k^*-k^*)} + \left[\frac{\sigma^2}{2} + 2\rho + \frac{1}{4}(k^*+k^* - 2b)^2\right] = \frac{b-k^*}{f(0)} - \frac{b-k^*}{f(0)} = 0$ , hence  $U_1'(k) < 0$ .

c) On the interval  $k \in \left(b, 2b-k^* + 2\sqrt{B - \frac{5}{2}\sigma^2}\right]$  we also have  $\frac{\sigma^2}{2} + 2\rho + \frac{1}{4}(k+k^* - 2b)^2 \leq 0$ . Then  $U_1'(k) = -2(k-b)[1-F(k-k^*)] + \left[\frac{\sigma^2}{2} + 2\rho + \frac{1}{4}(k+k^* - 2b)^2\right] f(k-k^*) < 0$ .

d) On the interval  $k > 2b-k^* + 2\sqrt{B - \frac{5}{2}\sigma^2}$  we also have  $k+k^* - 2b > 0$ . Hence, on this interval  $g''(k) > 0$ . Also notice that  $g(2b-k^* + 2\sqrt{B - \frac{5}{2}\sigma^2}) < 0$ . Then two cases are possible: (i)  $g'(2b-k^* + 2\sqrt{B - \frac{5}{2}\sigma^2}) \geq 0$ . Then on the whole interval  $g'(k) > 0$  and  $g(k)$  is increasing. As  $g(2b-k^* + 2\sqrt{B - \frac{5}{2}\sigma^2}) < 0$  and  $g(+\infty) > 0$ , there exists  $k_0$ :  $g(k) < 0$  for  $k < k_0$  and  $g(k) > 0$  for  $k > k_0$ . Now  $U_1(k)$  is decreasing for  $k < k_0$  and increasing for  $k > k_0$ . Hence, to show that  $k^*$  is a global maximum it is enough to check that  $U_1(k^*) \geq U_1(+\infty) = 0$ , which was already done.

(ii)  $g'(2b-k^* + 2\sqrt{B - \frac{5}{2}\sigma^2}) < 0$ . Then there exists  $k^{**} > 2b-k^* + 2\sqrt{B - \frac{5}{2}\sigma^2}$ : for  $k < k^{**}$   $g(k)$  is decreasing; for  $k > k^{**}$   $g(k)$  is increasing. As  $g(+\infty) > 0$ , there exists  $k_0$ :  $g(k) < 0$  for  $k < k_0$  and  $g(k) > 0$  for  $k > k_0$ . Then  $U_1(k)$  is decreasing for  $k < k_0$  and increasing for  $k > k_0$  and  $k^*$  is a global maximum as  $U_1(k^*) \geq U_1(+\infty) = 0$ .

<sup>37</sup>Recall that  $v(k-k^*) = \frac{f(k-k^*)}{1-F(k-k^*)}$ .

We now verify that  $k_U = k^* > k_U^{bas.}$ :

$$\begin{aligned}
k_U &= b + \left( \sqrt{\pi} - \sqrt{\pi + \frac{B}{\sigma^2} - \frac{5}{2}} \right) \sigma > k_U^{bas.} = b + \left( 1 - \frac{B}{2\sigma^2} \right) \frac{\sigma}{\sqrt{\pi}} \\
\iff \sqrt{\pi} - \sqrt{\pi + \frac{B}{\sigma^2} - \frac{5}{2}} &> \left( 1 - \frac{B}{2\sigma^2} \right) \frac{1}{\sqrt{\pi}} \iff \sqrt{\pi} + \frac{1}{\sqrt{\pi}} \left( \frac{B}{2\sigma^2} - 1 \right) > \sqrt{\pi + \frac{B}{\sigma^2} - \frac{5}{2}} \\
\iff \pi + \frac{B}{\sigma^2} - 2 + \frac{1}{\pi} \left( \frac{B}{2\sigma^2} - 1 \right)^2 &> \pi + \frac{B}{\sigma^2} - \frac{5}{2} \iff \frac{1}{2} + \left( \frac{B}{2\sigma^2} - 1 \right)^2 > 0.
\end{aligned}$$

□

*Proof of Proposition 13.* Start with calculation of marginal utilities:

$$U'_i(k_i) = -2(k_i - b_i)F(k_i - k_j) - \left[ \frac{\sigma^2}{2} + 2\rho + \frac{1}{4}(k_i + k_j - 2b)^2 \right] f(k_i - k_j)$$

Consider the symmetric case:  $k_1 = k_2 = k$ . The FOCs give two critical points:

$$k = b - \sqrt{\pi}\sigma - \sqrt{\left(\pi - \frac{5}{2}\right)\sigma^2 + B} \text{ and } k = b - \sqrt{\pi}\sigma + \sqrt{\left(\pi - \frac{5}{2}\right)\sigma^2 + B}.$$

Next, calculate second derivatives:

$$\begin{aligned}
U''_i(k_i) &= -2F(k_i - k_j) \\
&- \left[ 2(k_i - b_i) + \frac{1}{2}(k_i + k_j - 2b_i) + \left( \frac{B}{2\sigma^2} - \frac{5}{4} \right) (k_i - k_j) - \frac{1}{8\sigma^2}(k_i - k_j)(k_i + k_j - 2b_i)^2 \right] f(k_i - k_j)
\end{aligned}$$

We get that only  $k^* = b - \sqrt{\pi}\sigma + \sqrt{\left(\pi - \frac{5}{2}\right)\sigma^2 + B}$  satisfies SOCs.

In order to satisfy principal optimality we need  $k^* \leq 0$  or, equivalently, both  $\frac{b}{\sigma} \leq \sqrt{\pi}$  and  $B \leq b^2 - \sqrt{\pi}b\sigma + \frac{5}{2}\sigma^2$ .

Calculating, we get that  $U_i(k^*, k^*, H) = \frac{(\pi-1)\sigma}{\sqrt{\pi}} \sqrt{\left(\pi - \frac{5}{2}\right)\sigma^2 + B} - \left(\pi - \frac{7}{4}\right)\sigma^2$  (the same as in upward equilibrium). As in upward equilibrium case, a necessary condition is  $B \geq \left(\frac{5}{2} - \frac{3\pi(8\pi-11)}{16(\pi-1)^2}\right)\sigma^2$ .

From previous arguments downward equilibrium may exist only if  $B \in \left[\left(\frac{5}{2} - \frac{3\pi(8\pi-11)}{16(\pi-1)^2}\right)\sigma^2, b^2 - 2\sqrt{\pi}b\sigma + \frac{5}{2}\sigma^2\right]$  and  $\frac{b}{\sigma} \leq \sqrt{\pi}$ . This interval is non-empty if and only if  $\frac{b}{\sigma} \leq \frac{3\sqrt{\pi}}{4(\pi-1)}$ . Note also that  $B \leq b^2 - 2\sqrt{\pi}b\sigma + \frac{5}{2}\sigma^2 \leq \frac{5}{2}\sigma^2$ .

To finish the proof we show that if  $B$  lies on this interval, then  $k = k^*$  is not only a local, but also a global maximum of  $U_1(k, k^*, L)$ .

Denote  $r(k) = -2(k - b) - \left[ \frac{\sigma^2}{2} + 2\rho + \frac{1}{4}(k + k^* - 2b)^2 \right] w(k - k^*)$  (remind that  $w(k - k^*) = \frac{f(k - k^*)}{F(k - k^*)}$ ).

Then  $U'_1(k) = -2(k - b)F(k - k^*) - \left[ \frac{\sigma^2}{2} + 2\rho + \frac{1}{4}(k + k^* - 2b)^2 \right] f(k - k^*) = F(k - k^*)r(k)$  and  $sign(U'_1(k)) = sign(r(k))$ .

First and second derivatives of  $r(k)$  are:

$$r'(k) = -2 - \left[ \frac{\sigma^2}{2} + 2\rho + \frac{1}{4}(k + k^* - 2b)^2 \right] w'(k - k^*) - \frac{1}{2}(k + k^* - 2b)w(k - k^*);$$

$$r''(k) = - \left[ \frac{\sigma^2}{2} + 2\rho + \frac{1}{4}(k + k^* - 2b)^2 \right] w''(k - k^*) - (k + k^* - 2b)w'(k - k^*) - \frac{1}{2}w(k - k^*).$$

Notice that as  $\frac{b}{\sigma} \leq \sqrt{\pi}$ ,  $B \leq b^2 - \sqrt{\pi}b\sigma + \frac{5}{2}\sigma^2 \leq \frac{5}{2}\sigma^2$  and  $k^* = b - \sqrt{\pi}\sigma + \sqrt{(\pi - \frac{5}{2})\sigma^2 + B} \leq b$ .

a) On interval  $k > b$   $U'(k) < 0$ , so there is no candidate for maximum there.

b) On interval  $k < b$  we also have  $k + k^* - 2b \leq 0$ . As  $w > 0$ ,  $w' < 0$ ,  $w'' > 0$ , we have  $r''(k) < 0$ .

As  $r'(-\infty) > 0$  and  $r'(k^*) < 0$ , there exists  $k^{**} < k^* < b$ :  $r(k)$  is increasing for  $k < k^{**}$ ,  $r(k)$  is decreasing for  $k > k^{**}$ . As also  $r(-\infty) < 0$ ,  $r(k^* - 0) > 0$  and  $r(k^* + 0) < 0$ , there exists  $k_0$ :  $r(k) > 0$  only on  $(k_0, k^*)$ . Therefore,  $U_1(k)$  is decreasing on  $k < k_0$ , increasing on  $(k_0, k^*)$ , decreasing on  $(k^*, b)$ . Hence,  $k^*$  is a global maximum if  $U_1(k^*) \geq U_1(-\infty) = 0$ , which has already been shown.  $\square$

## References

- Alonso, R. and N. Matouschek (2008). Optimal delegation. *The Review of Economic Studies* 75, 259–293.
- Ambrus, A. and S. Lu (2014). Almost fully revealing cheap talk with imperfectly informed senders. *Games and Economic Behavior* 88, 174–189.
- Ambrus, A. and S. Takahashi (2008). Multi-sender cheap talk with restricted state spaces. *Theoretical Economics* 3, 1–27.
- Austen-Smith, D. (1993). Interested experts and policy advice: multiple referrals under open rule. *Games and Economic Behavior* 5, 3–43.
- Battaglini, M. (2002). Multiple referrals and multidimensional cheap talk. *Econometrica* 70, 1379–1401.
- Battaglini, M. (2004). Policy advice with imperfectly informed experts. *Advances in Theoretical Economics* 4, 1–32.
- Bernhardt, D., M. Campello, and E. Kutsoati (2006). Who herds? *Journal of Financial Economics* 80, 657–675.
- Bhattacharya, S. and A. Mukherjee (2013). Strategic information revelation when experts compete to influence. *The RAND Journal of Economics* 44, 522–544.
- Billingsley, P. (2009). *Convergence of probability measures*, Volume 493. John Wiley & Sons.
- Calvert, R. L. (1985). Robustness of the multidimensional voting model: Candidate motivations, uncertainty, and convergence. *American Journal of Political Science* 29, 69–95.
- Chen, Q. and W. Jiang (2006). Analysts' weighting of private and public information. *Review of Financial Studies* 19, 319–355.
- Crawford, V. and J. Sobel (1982). Strategic information transmission. *Econometrica: Journal of the Econometric Society* 50, 1431–1451.

- Dessein, W. (2002). Authority and communication in organizations. *The Review of Economic Studies* 69, 811–838.
- Dewatripont, M. and J. Tirole (1999). Advocates. *Journal of Political Economy* 107, 1–39.
- Downs, A. (1957). *An economic theory of democracy*. New York: Harper & Row.
- Esponda, I. (2008). Behavioral equilibrium in economies with adverse selection. *The American Economic Review* 98, 1269–1291.
- Eyster, E. and M. Rabin (2005). Cursed equilibrium. *Econometrica* 73, 1623–1672.
- Friedman, D. (1991). A simple testable model of double auction markets. *Journal of Economic Behavior and Organization* 15, 47–70.
- Gentzkow, M. and E. Kamenica (2015). Competition in persuasion. Working paper.
- Gerardi, D., R. McLean, and A. Postlewaite (2009). Aggregation of expert opinions. *Games and Economic Behavior* 65, 339–371.
- Gilligan, T. and K. Krehbiel (1989). Asymmetric information and legislative rules with a heterogeneous committee. *American Journal of Political Science* 33, 459–490.
- Holmström, B. (1977). *On incentives and control in organizations*. Stanford University.
- Hotelling, H. (1929). Stability in competition. *The Economic Journal* 39, 41–57.
- Iezzoni, L., S. Rao, C. DesRoches, C. Vogeli, and E. Campbell (2012). Survey shows that at least some physicians are not always open or honest with patients. *Health Affairs* 31, 383–391.
- Jehiel, P. (2005). Analogy-based expectation equilibrium. *Journal of Economic theory* 123, 81–104.
- Jehiel, P. and F. Koessler (2008). Revisiting games of incomplete information with analogy-based expectations. *Games and Economic Behavior* 62, 533–557.
- Kamenica, E. (2008). Contextual inference in markets: On the informational content of product lines. *The American Economic Review* 98, 2127–2149.
- Kartik, N., F. Squintani, and K. Tinn (2015). Information revelation and pandering in elections. Technical report, mimeo, Columbia University.
- Klemperer, P. (1999). Auction theory: a guide to the literature. *Journal of Economic Surveys* 13, 227–286.
- Krishna, V. and J. Morgan (2001a). Asymmetric information and legislative rules: Some amendments. *American Political Science Review* 95, 435–452.
- Krishna, V. and J. Morgan (2001b). A model of expertise. *Quarterly Journal of Economics* 116, 747–775.

- Krishna, V. and J. Morgan (2008). Contracting for information under imperfect commitment. *The RAND Journal of Economics* 39, 905–925.
- Li, H. and W. Suen (2004). Delegating decisions to experts. *Journal of Political Economy* 112, 311–335.
- Melumad, N. and T. Shibano (1991). Communication in settings with no transfers. *The RAND Journal of Economics* 22, 173–198.
- Milgrom, P. and J. Roberts (1986). Relying on the information of interested parties. *The RAND Journal of Economics* 17, 18–32.
- Milgrom, P. and R. Weber (1985). Distributional strategies for games with incomplete information. *Mathematics of Operations Research* 10, 619–632.
- Morris, S. and H. Shin (2002). The social value of private information. *The American Economic Review* 92, 1521–1534.
- Morris, S. and H. Shin (2003). Global games: Theory and applications. Econometrics Society Monographs.
- Myatt, D. and C. Wallace (2014). Central bank communication design in a Lucas-Phelps economy. *Journal of Monetary Economics* 63, 64–79.
- Ottaviani, M. and P. Sørensen (2006). Professional advice. *Journal of Economic Theory* 126, 120–142.
- Pesendorfer, W. and A. Wolinsky (2003). Second opinions and price competition: Inefficiency in the market for expert advice. *The Review of Economic Studies* 70, 417–437.
- Pétry, F. and B. Collette (2009). Measuring how political parties keep their promises: A positive perspective from political science. Volume 15 of *Studies in Public Choice*, pp. 65–80.
- Reed, M. and B. Simon (1980). *Methods of modern mathematical physics: Functional analysis*, Volume 1. Gulf Professional Publishing.
- Royden, H. and P. Fitzpatrick (1988). *Real analysis*, Volume 4. Prentice Hall New York.
- Sampford, M. (1953). Some inequalities on Mill’s ratio and related functions. *The Annals of Mathematical Statistics* 24, 130–132.
- Shin, H. (1998). Adversarial and inquisitorial procedures in arbitration. *The RAND Journal of Economics* 29, 378–405.
- Treitel, G. (1999). *The law of contract* (10 ed.). Sweet & Maxwell.
- Wittman, D. (1983). Candidate motivation: A synthesis of alternative theories. *American Political Science Review* 77, 142–157.

Wolinsky, A. (2002). Eliciting information from multiple experts. *Games and Economic Behavior* 41, 141–160.

Zitzewitz, E. (2001). Measuring herding and exaggeration by equity analysts. Technical report, mimeo, Stanford University.